# JOURNAL OF APPLIED PHYSICS

# JOURNAL OF APPLIED PHYSICS

Vol. 20, No. 2 FEBRUARY, 1949


## *In This Issue*


*Page*

**Photo-Density Method of Determining the Relative Penetration of Diffused Sodium 24 Tracer into Glass** — James R. Johnson 129

**Measurement of Forces Resisting Armor Penetration** — A. Victor Masket 132

**Probability Criterion for the Design of Servomechanisms** — John R. Ragazzini and Lotfi A. Zadeh 141

**Measurements of the Mechanical Properties of Polymer Solutions by Electromagnetic Transducers** — Thor L. Smith, John D. Ferry, and Frederic W. Schremp 144

**Density and Packing in an Aggregate of Mixed Spheres** — Douglas Rennie Hudson 154

**Resistance of Oxide Cathode Coatings for High Values of Pulsed Emission** — W. E. Danforth and D. L. Goldwater 163

**Particle Size Determination by Soft X-Ray Scattering** — K. L. Yudowitch 174

**Aberration Correction with Electron Mirrors** — E. G. Ramberg 183

**Sweep Frequency Ionosphere Equipment** — Peter G. Sulzer 187

**Work Functions and Conductivity of Oxide-Coated Cathodes** — G. W. Mahlman 197

**Positive-Replica Technique for Electron Microscopy** — C. M. Schwartz, A. E. Austin, and P. M. Weber 202

**Mathematical Theory of Photo-Viscoelasticity** — Raymond D. Mindlin 206

**Air Friction on Rapidly Moving Surfaces** — A. R. Kuhlthau 217

**Letters to the Editor:**

**On the Flow of a Perfect Fluid between Two Plates from a Circular Inlet Channel** — W. R. Smythe 224

**Small Spherical Particles of Exceptionally Uniform Size** — Robert C. Backus and Robley C. Williams 224

**Coefficient of Anomalous Viscosity** — Charles Mack 225

**Measurement of Gain of Electromagnetic Horns** — Allen S. Dunbar and M. Donald Adcock 226

**Method of Measuring Dynamic Mechanical Properties of Rubber-Like Materials** — K. W. Hillier, H. Kolsky, and J. P. Lewis 227

**Comment on "A Method of Measuring Dynamic Mechanical Properties of Rubber-Like Materials"** — A. W. Nolle 227

**Extension of MacFarlane's Method of Deducing Refractive Index from Radio Observations** — A. W. Straiton 228

**Errata: Energy Losses of Sound Waves in Metals due to Scattering and Diffusion** — W. P. Mason and H. J. McSkimin 228

**Note on the Ionospheric Absorption Problem** — Leslie G. McCracken 229

**Grain Growth in Octachloropropane** — W. C. McCrone and P. T. Cheng 230

**Comments on "Grain Growth in Octochloropropane"** — Paul A. Beck 231

**Comments on "Electronic Radiography and Microradiography"** — Herman E. Seemann 231

*In Every Issue*

**Calendar of Meetings** 216

**New Booklets** 232

**Here and There** 233



**BOARD OF EDITORS**

Elmer Hutchisson, *Editor*

*Associate Editors*

R. Bowling Barnes
Otto Beeck
W. James Lyons
Philip Morrison
Paul E. Sabine
Raymond J. Seeger
Frederick Seitz
Charles S. Smith
J. R. Townsend

**Publication of the American Institute of Physics**

**OFFICERS**

George R. Harrison, *Chairman*
Wallace Waterfall, *Secretary*
George B. Pegram, *Treasurer*
John T. Tate, *Adviser on Publications*

**GOVERNING BOARD**

Luis W. Alvarez
Robert B. Brode
E. U. Condon
A. J. Dempster
W. F. Fair, Jr.
Harvey Fletcher
R. C. Gibbs
George R. Harrison
W. V. Houston
Paul E. Klopsteg
F. W. Loomis
Philip M. Morse
George B. Pegram
A. H. Pfund (*deceased*)
John T. Tate
Wallace Waterfall
M. W. White
A. G. Worthing

**Consultant on Printing**

Melvin Loos

**ADMINISTRATIVE STAFF**

Henry A. Barton, *Director*
Cleveland Norcross, *General Manager*
Theodore Vorburger, *Business Manager*
Ruth F. Bryans, *Publication Manager*
Betty Hofstadter, *Office Manager*
Eileen B. Neuberger, *Circulation Manager*



The Journal of Applied Physics, published monthly at Prince and Lemon Sts., Lancaster, Pa., is devoted to physics in its role as the science basic to other natural sciences and to the arts and industries. Previous to 1937 this publication was known as Physics. The Journal of Applied Physics publishes editorials, news of physicists, and reviews, as well as technical papers of applied physics. Articles appearing in it are indexed regularly, according to their field, by Physics Abstracts, Chemical Abstracts, Engineering Index, Inc., Industrial Arts Index, etc.

**Manuscripts** should be submitted to Elmer Hutchisson, Editor, Case Institute of Technology, Cleveland 6, Ohio. The authors' institutions are requested to pay a **publication charge** of $4 per page which, if honored, entitles them to 100 free reprints. Instructions will be sent with galley proofs.

**Subscriptions** and orders for back numbers should be addressed to Prince and Lemon Streets, Lancaster, Pa. or to the American Institute of Physics, 57 East 55 Street, New York 22, N. Y.

**Subscription Price**

| | U.S. and Canada | Elsewhere |
|---|---|---|
| To members and associate members of the American Institute of Physics | $7.00 | $ 8.00 |
| To all others | 9.00 | 10.00 |

**Back Number Prices**

Complete set: Vol. 1, July, 1931–Vol. 19, 1948—$140.00
Yearly back number rate: $10.00
Single copies: $1.00 each.

**Changes of address** should be sent to the Business Manager, American Institute of Physics; **corrected proofs** should be sent to the Publication Manager, American Institute of Physics.

**Advertising rates** supplied on request. Orders, advertising copy, and cuts should be sent to the American Institute of Physics.

Entered as second class matter January 22, 1937, at the Post Office at Lancaster, Pennsylvania, under the Act of March 3, 1879.
Accepted for mailing at the special rate of postage provided for in the Act of February 28, 1925, authorized May 2, 1932.



Copyright 1949, by the American Institute of Physics.

Journal
of
Applied Physics

Volume 20, Number 2 February, 1949

# Photo-Density Method of Determining the Relative Penetration of Diffused Sodium 24 Tracer into Glass

JAMES R. JOHNSON
*Department of Ceramic Engineering, The Ohio State University, Columbus, Ohio*
(Received August 5, 1948)

A photo-density method for the determination of diffusion of radioactive tracer is presented and compared in detail to a thin section analysis method. The method is no doubt applicable to similar tracer problems such as metal diffusion studies. The proper analysis of an emulsion exposed by decaying $Na^{24}$ which had been allowed to diffuse into glass is given. The necessity for making and using an emulsion gamma-curve for the tracer in this analysis is shown.

IN the study of the diffusion of $Na^{24}$ through glass, a fast and reliable technique was required. The first successful method developed for this investigation was a technique of removing successive thin layers of glass by grinding the sample with abrasive on microscope slides. The slides were then measured for relative activity by means of a Geiger counter. This method was sufficiently reliable and accurate but lacked the quality of being fast.

It was believed that a photographic method might succeed if it were properly carried out. Certainly it would simplify the experimental work on this and similar investigations if it were possible to analyze the penetration of the tracer merely by placing the sample on a photosensitive emulsion and allowing it to make the desired record. The difficulty proved to be the proper analysis of the exposed emulsion.

It is well known that the sensitivity of an emulsion is not directly proportional to the intensity of the impinging radiation. Therefore, for any emulsion analysis, a gamma-curve must be constructed. Such a curve is a plot of the log of the relative impinging radiation *vs.* the density of the corresponding exposed emulsion. Since the shape and slope of this curve varies with different wave-lengths of radiation, it seemed certain that it would be necessary to construct a curve specifically for the beta-emission of $Na^{24}$ in this case. This was done and the resulting gamma-curve provided a tool for analyzing the photographic penetration record of the $Na^{24}$ tracer.

The specimen containing the diffused $Na^{24}$ was placed directly on the emulsion for one-half hour. Figure 5 illustrates the process schematically. The plate was then developed and analyzed in a photo-densitometer. By properly correlating the density-penetration curve thus obtained (Fig. 4) with the gamma-curve, it was possible to obtain relative concentration-penetration values. The actual end product was a plot of the log relative concentration *vs.* the square of the depth of penetration, a straight line slope (Fig. 1). It is to be noted that only the slope of the line is significant. A comparison of the slope obtained by the section grinding method with that obtained by the photo method shows that the two are

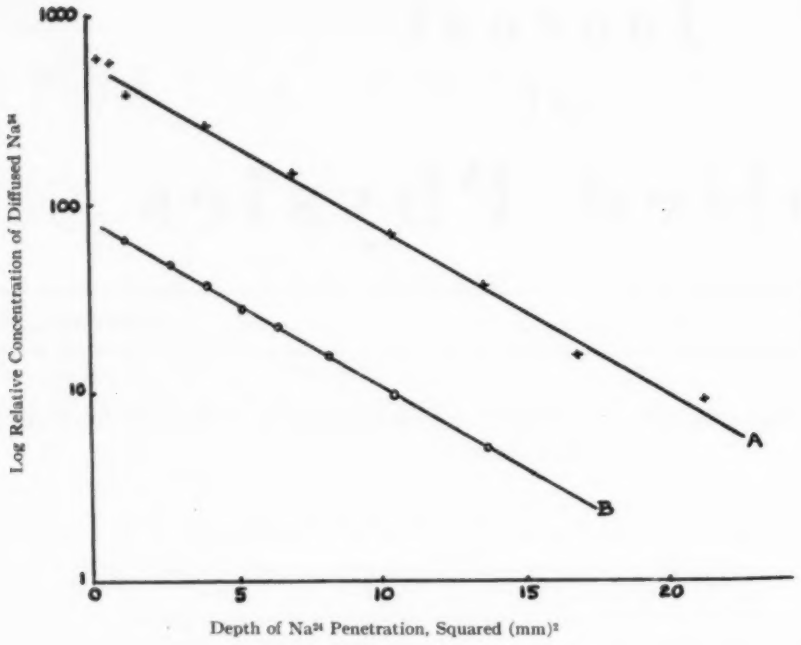



FIG. 1. Penetration of glass by $Na^{24}$ at 1250°F. Curve A was obtained by grinding method; B by photo method.

identical. By way of a check the work has been repeated with similar results each time. The photo method has given better series of points in every case.

The photo-density method has many advantages other than the obvious one of saving time. It is probably less subject to error and it eliminates the laborious processes of the grinding method such as grinding, weighing and measuring as well as the counting procedure with the Geiger counter. It makes possible the investigation of several specimens at a time.

The emulsion used in this work was E.K. SA No. 1 on 4×8 plates. It must be emphasized that careful control of developing time and temperature is necessary for successful results. These two factors must be the same for each plate. It is the author's opinion that the method could be applied to other radioactive tracers, which have sufficient activity for exposing an emulsion, by constructing a gamma-curve for the emission from the tracer concerned and applying the foregoing method of analysis to the experimental photographic record using that tracer.

The procedure for investigating any one tracer such as $Na^{24}$ involves the calculation of the proper exposure times for making a gamma-curve, the plotting of the gamma-curve, and the analysis of

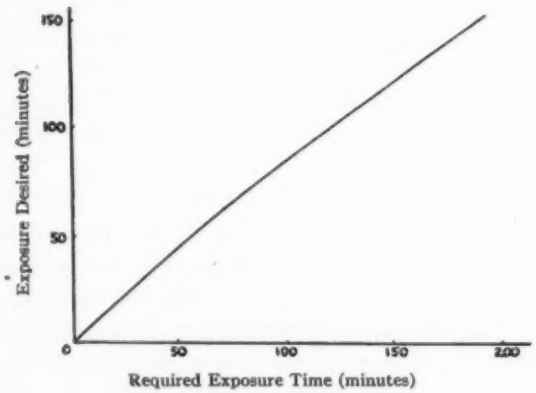



FIG. 2. $Na^{24}$ exposure chart.

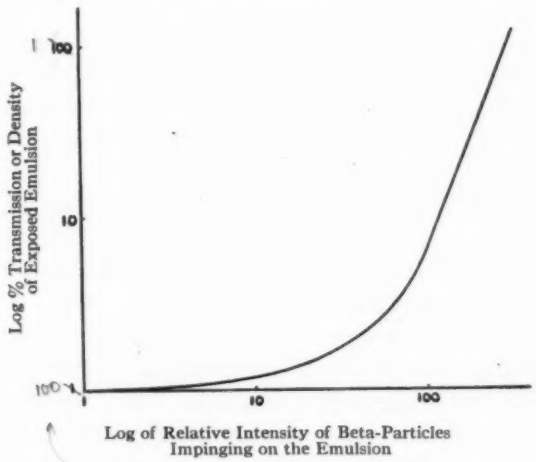



FIG. 3. Gamma-curve for $Na^{24}$.

the tracer photograph using that gamma-curve. The equation used for calculation of exposure times is

$$t = A \cdot (N/I)$$

where $t$ is the exposure time, $A$ is a constant, $N$ is the number of atoms decaying in that time and $I$ is the rate of decay during that time. This equation is an approximation and would not hold for large values of $t$ or for elements with very short half-lives. In either of these cases, the original equation is applicable,

$$t = -(1/\lambda)\ \ln[1-(N\lambda/I)]$$

or

$$t = (1/\lambda)[x+(x^2/2)+(x^3/3)\cdots]$$

where

$$x = N\lambda/I.$$

Since only relative values are desired for the first calculation, $I_0$ may be given an approximate value, and $N$ and $A$ can be chosen such that $t$ equals some convenient time such as 100 seconds. Subsequently, $N_2$ is chosen to equal $2\times N_1$. $N_3$ will equal $2\times N_2$ etc. and corresponding $I_0$ values will depend on the previously accumulated time and the time of the ensuing exposure and can be determined from a decay curve. Figure 2 shows a curve set up for use with $Na^{24}$ indicating proper exposure times for plotting a gamma-curve. The emulsion is then exposed successively for the proper values of $t$ and then developed at a specific temperature and time. The percent transmission of each of these exposures is plotted *vs.* the relative values of $N$ as shown in Fig. 3, the gamma-curve. A typical penetration-film density curve for the diffusion of $Na^{24}$ into glass is shown in Fig. 4. To get the proper value of the relative concentration of the diffused $Na^{24}$ from this curve at any point, it is necessary only to read it from the gamma-plot according to its corresponding percent transmission. Figure 5 illustrates the process schematically.

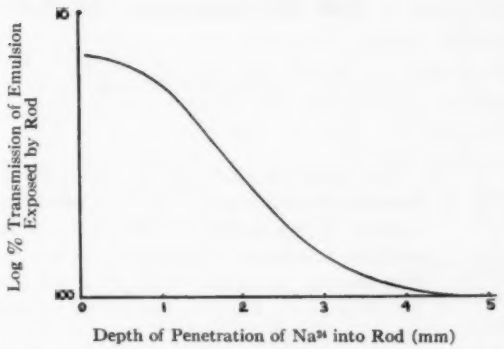



FIG. 4. Penetration curve of $Na^{24}$ into glass at 1250°F.

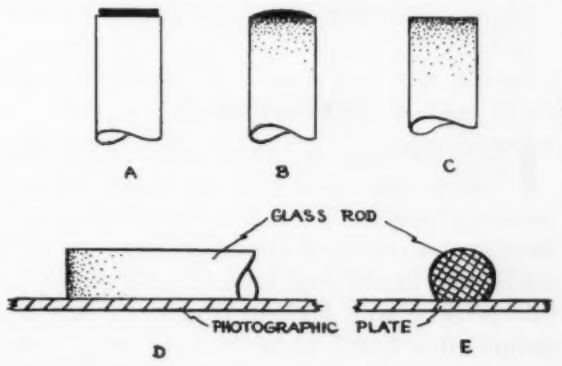



FIG. 5. Schematic representation of experimental technique. A shows the glass rod with radioactive sodium carbonate in place before firing. B shows the rod after removal from the furnace. C shows it with the remaining sodium carbonate removed, ready for grinding or emulsion exposure. D and E show the rod as it is placed in the exposure process. The dots represent diffused $Na^{24}$.


## ACKNOWLEDGMENT

The author wishes to express his thanks for the suggestions, assistance, and use of equipment in developing this technique to Dr. M. L. Pool, Dr. H. H. Blau, Dr. Q. Van Winkle, N. N. Ault, and D. Timma.

# The Measurement of Forces Resisting Armor Penetration*

A. Victor Masket**
*Mechanics Division, Naval Research Laboratory, Washington, D. C.*
(Received June 14, 1948)

This paper summarizes the experimental and theoretical status of the optical chronograph developed in the course of ballistic research at the Naval Research Laboratory. The instrument, together with a simple procedure for analysis of data, is capable of yielding the position, velocity, and deceleration of a non-plastically deforming small arms projectile during the armor penetration process, which lasts from 30 to 150 microseconds. Decelerations have been observed up to $5\times10^7$ ft./sec.$^2$. The longitudinal vibrations induced in projectiles during impact were observed by a shadow-photography technique which yields time-displacement curves from the motion of the trailing face of a projectile during penetration. The precision of the derived decelerations, about 4 percent, is sufficient to permit the evaluation of strain-rate and inertia effects during high speed indentation by means of conical indenters at strain-rates approaching $2\times10^5$/sec.

## INTRODUCTION

THE wide interest which arose in the problems associated with the development of improved armor plate resulted in an attempt to formulate a theory of armor penetration. Up to the Second World War, theoretical knowledge on this subject was limited largely to empirical formulations based on laboratory firings, armor proving tests, or upon rough data on the dynamic behavior of the armor or projectile *during the penetration cycle.* This lack of accurate data was a consequence of the difficult experimental problem of continuously observing the penetration of an armor specimen by a non-deforming bullet. It is the object of this paper (1) to present a description of a precision optical chronograph utilizing a high speed drum camera technique for continuous observation of projectile motion during armor penetration, and (2) to describe the application of the chronograph to determine the forces which non-deforming bullets encounter while penetrating armor plate.

Prewar security classification prevented timely publication of earlier stages in the design and use of the optical chronograph carried on by G. R. Irwin, Frank Louckes, and G. D. Kinzer, who established the basic technique and reported force-time data for various target steels. The scatter of film measurements was then too great to resolve vibration effects clearly.

## APPARATUS

A strip of film placed in a high speed revolving drum affords a convenient mechanical technique for expanding the time scale in order to record events which occur during a short time interval such as the penetration of armor by a small calibre projectile. In such ballistic phenomena, the entire process is known to occupy about $10^{-4}$ seconds or less. As a consequence it would be of little use to attempt "shutter" photography unless the number of frames per second could be made to approach $10^6$. The continuous shadow-photography technique which is made possible by the apparatus described here allows the experimenter to determine the position, velocity, and deceleration of the trailing edge of the bullet at intervals of $5\times10^{-7}$ sec. under the best conditions now obtainable.

The physical principle of the apparatus consists of obtaining the shadow of a projectile on a strip of film moving perpendicular to the line of flight of the projectile. The line of flight is arranged so that the projectile progressively obstructs the illumination of a fine slit placed very close to the surface of the fast moving film. The schematic diagram of the apparatus is shown in Fig. 1. The bullet leaves the bore of the gun about 18″ from the armor specimen and, when about 12″ from the specimen, it completes an

* Published by permission of the U. S. Navy. (The opinions expressed in this paper are the author's and do not necessarily represent those of the U. S. Navy.) The work summarized in this paper is based upon the following declassified reports: NRL 0-1591, 0-2275, 0-2863 and was performed by initial authorization of Bureau of Ordnance 1tr. S13-1(4/173)(Q8) of 12/13/34.
** Now in the Department of Physics, University of North Carolina, Chapel Hill, North Carolina.

*RC* circuit by making contact with two metallic foil screens of negligible ballistic resistance. The resulting discharge of a capacitor through a resistor puts a positive voltage pulse on the grid of a biased thyratron which, on conduction, initiates the action of a variable time-delay circuit controlling the spark light-source. The spark is timed to occur at about the instant the projectile begins to obstruct the light path between the collimating and objective lenses. The flight path is made to pass through the apex of a wedge of light created by the cylindrical lens of the collimating system and is focussed on the slit which is parallel and very close to the rotating film.

The light source is produced by discharging a $5\frac{1}{2}$-$\mu$f capacitor, charged to 15,000 volts, through a $\frac{1}{2}$-mm bore glass capillary so masked that the light emerges from a pinhole aperture placed at the principle focus of the first collimating lens. The duration of the discharge has been extended to about 600 microseconds by placing a 75-ohm resistor in series with the capillary gap. A leader stroke preceding the main discharge by about 10 microseconds is easily observable on the films. No oscillation of the discharge has been encountered during the usable portion of light emission. Several advantages have arisen from the present arrangement of having the illuminated slit located in the focal plane of the camera assembly (suggested by G. D. Kinzer) rather than in the objective plane through the trajectory. When the slit was in the objective plane, the impact of the projectile produced "splash" from the surface of metallic armor specimens which marred the shadowgrams by streaking across the slit and damaged its precision ground edges by ricochet. These deleterious effects on the films and slit, observed in initial experiments, were eliminated by the relocation of the slit. In addition, the placing of the slit very close to the film surface eliminates the blurring due to diffraction effects. The slit width is still governed by intensity requirements. With the light intensity now available, a slit width of 0.002″ is permissible for a film speed of 500 ft./sec.

The air-driven turbine and hollow drum unit (Fig. 1) has been precision turned and machined from a single block of duralumin alloy. The drum will hold a foot-length strip of 35-mm film on its inner surface. The assembly of the drum is such that it may be moved independently of the lenses and mirrors of the optical system for ease of loading in darkness. The drum with film is positioned while turning at high speed so that the film, flattened against the drum by centrifugal force, will not be damaged by the masking slit which is located only 0.0025″ from the focal plane.

A photograph of the assembly appears in Fig. 2. Directly in front of the gun, a blast deflector (not shown in Fig. 1) obscures the armor specimen and its holder in the rear of the picture. The bank of high voltage condensers seen on the right discharges through the capillary gap mounted on the insulators. On the left is the camera assembly with the drum mounting raised into position. A photograph of the drum in its housing is given in Fig. 3. The design of the yoke holding an upper bearing for the rotating drum permits the placement of a mirror at 45° to the vertical (as shown in Fig. 1) to reflect the trajectory image normal to the film.

The instruments (not shown in Fig. 2) for control and measurement of drum speed, time delay of discharge, and capillary pressure are

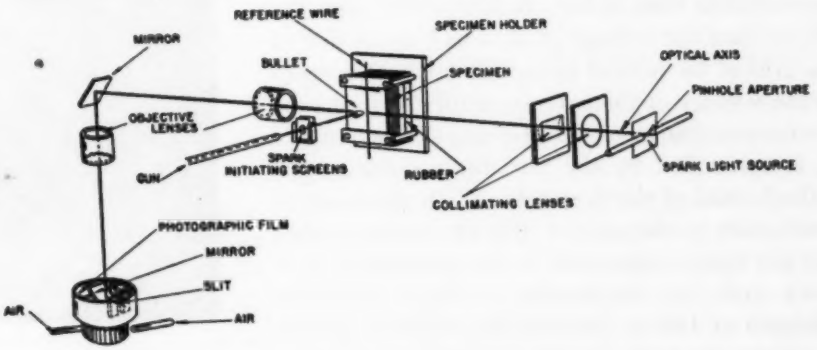



FIG. 1. Schematic diagram of apparatus.

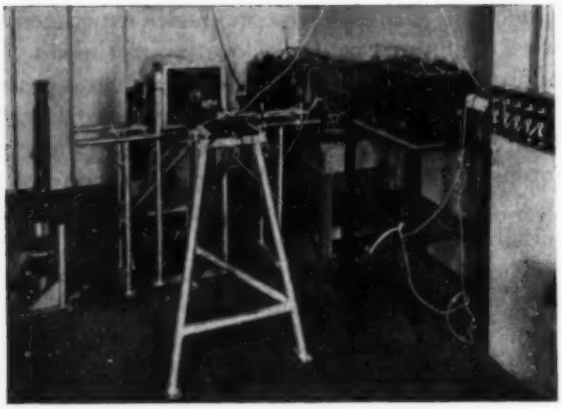


FIG. 2. Photograph of assembly.

located in another room and connect to the force-curve apparatus through the rear wall.

Since the drum speed is measured only for intervals less than 0.1 millisecond, the drum inertia keeps the speed sufficiently constant when the speed is regulated with a conventional air-pressure reduction valve.

Drum speed is measured by matching the signal from a magnetic pick-up from the turbine driving the drum with the signal from a calibrated variable frequency generator, General Radio Company Interpolator Oscillator, Type 617-C. The magnetic pick-up signal is produced by small steel inserts in the turbine which vary the reluctance in a magnetic circuit passing through the turbine. The gun is fired when the matching is viewed as a stationary Lissajous figure on an oscilloscope screen.

The time-delay circuit is outlined in Fig. 4. Plate voltage was obtained from batteries since they were good for their shelf life and filament heater supply was obtained from a 6-v transformer. Constant plate voltage is needed for reproducible time delay. It is evident from the circuit that the voltage at $A$, which is applied to the grid of the second tube, will rise with respect to the voltage of the cathode while the first tube discharges condenser $C_2$. The rate of discharge of $C_2$ is controlled by the variable resistor in the cathode lead of the first tube. With the onset of conduction in the second 2050 thyratron, a high voltage surge, generated in the secondary of a spark coil by discharging a 10-$\mu$f capacitor (charged at 180 v) through the primary, is used to trigger the spark for the light source.

## APPLICATION

Several stages during the shadow photography of a bullet approaching and penetrating an armor specimen are depicted in Fig. 5. In stage $A$ the shadow image $P$ of the bullet is shown at an instant during its approach to the shadow image of the armor, while the film is moving left, normal to the image of the line of flight. The film is shown blackened, by exposure to light through the masking slit, over those areas which are not shielded by the shadow image of the bullet. The shadow of the bullet appears as an unexposed portion of the film bounded by inclined parallel edges. The tangent of their inclination to the direction of motion of the film is proportional to the speed of the bullet for a given film speed. Two reference marks which appear on each shadowgram are essential for film analysis. These marks are produced by shadows of two very fine wires. One wire is placed parallel to the armor plate at a measured distance from it and perpendicular to the line of flight. The second wire is fixed across the masking slit. The two shadow streaks are shown at $a$ and $b$ in Fig. 5. They accurately determine the direction of motion of the film and therefore of the bullet and, at the same time, determine the position of the base of the bullet at the in-

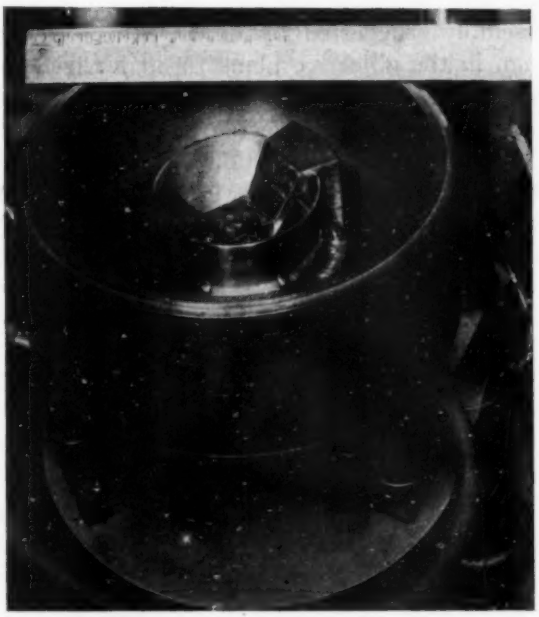


FIG. 3. Photograph of drum in housing.

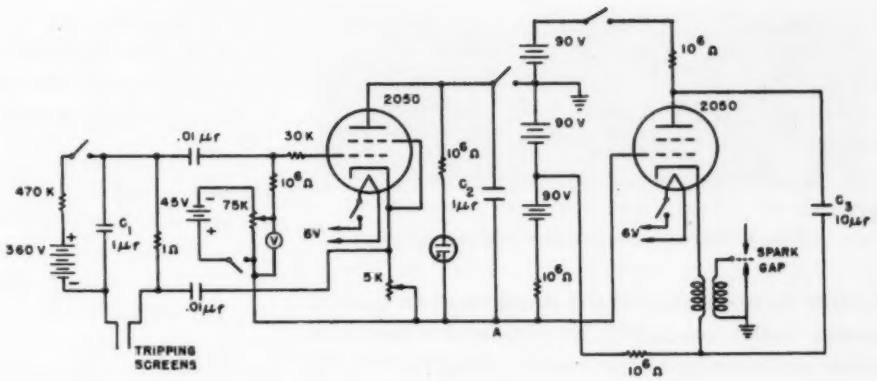



FIG. 4. Time delay circuit.

stant its nose strikes the armor. This instant is illustrated in stage *B*. It is clear that the base of the bullet is at a distance *d* below the shadow *a* and *d* is given by the length of the bullet less the distance *c*.

The appearance of the exposed areas of the film, up to the instant a bullet has been completely brought to rest by the armor specimen, is shown in stage *C*. The edge of the shadowgram corresponding to the motion of the bullet base is curved, of course, during impact and is a time-displacement curve from which the essential data of the penetration is obtained. (It is to be noted, however, that the curve so obtained is not the same as that corresponding to the motion of the center of mass of the bullet because the exceedingly large forces encountered by bullets will introduce longitudinal vibrations which are observable by the technique described.) Should stage *C* also represent the possibility that the bullet remains imbedded in the armor, then the area shown in cross-hatching would subsequently become exposed and darkened to the extent that the discharge persisted.

An illustrative shadowgram obtained from the force-curve apparatus is shown in Fig. 6. This unusual shadowgram shows not only the approach and impact of a bullet but also the direct rebound along the line of flight from the armor specimen. While not visible in the enlargement, the original film shows clearly the sinusoidal ripple along the edge of the shadowgram during the rebound (when the bullet is in free flight) which is a consequence of the periodic variation of the base velocity with respect to the velocity of the center of mass while the bullet is in longitudinal vibration from the impact.

The primary data in the film analysis consists of a tabulated set of slope measurements at equal intervals of time before and during impact along the edge of the shadowgram made by the base of the bullet. These slope measurements were made by judging the tangency of a cross hair whose consequent inclination to the direction of film motion was read from a goniometer attachment on a Gaertner comparator. At first, straightforward differences of slope values, corrected for momentum transfer to the plate, were smoothed to give values proportional to the deceleration of the bullet. A typical group of such curves for identical bullets fired against several thicknesses of armor is given in Fig. 7. These show the periodic secondary maxima when the deceleration of the base is not corrected to account for longitudinal vibration. In particular one notes that the trailing edge of the bullet does not show any deceleration until a half period after impact has begun. This is just the time needed for the compressional wave to travel from the nose to the base. The deceleration of the center of mass can be obtained very simply, however, from the primary data and yields the actual force on the nose of the bullet during the penetration into armor.

## THEORETICAL TREATMENT OF DATA

In order to treat the longitudinal vibrations of a projectile analytically, the following assumptions were made: (1) that if the pointed end of the projectile is short in comparison with its total length, it may be considered as a cylinder of equivalent mass and cross section but calculably shorter in length, and (2) that in normal impact, transverse vibrations are negligible.

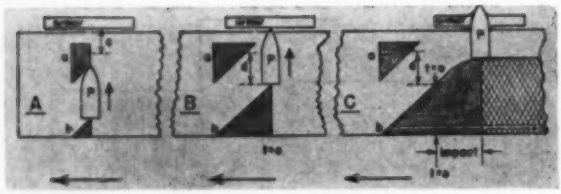



FIG. 5. Stages during shadow photography.

Under these conditions the problem of an ogival-nosed bullet executing longitudinal vibrations while penetrating armor can be simplified to the problem of an elastic rod moving with velocity $V_0$ which at time $t=0$ encounters a force, $F(t)$, see Fig. 8, acting on the leading face and opposing the motion. For small strains and damping, the equation to be solved is

$$\partial^2 u(x,t)/\partial t^2 = c^2(\partial^2 u(x,t)/\partial x^2), \quad c^2 = E/\rho, \quad (1)$$

subject to the boundary conditions:

$$t=0\colon\ u\equiv 0,\ \partial u/\partial t|_{t=0} \equiv V_0; \quad (1a)$$

$$x=0\colon\ \partial u/\partial x|_{x=0} \equiv 0;$$

$$x=l\colon\ \partial u/\partial x|_{x=l} = F(t)/AE. \quad (1b)$$

The symbols have the following definitions:

$t$ = time, $x$ = position in projectile of length $l$ relative to the base ($x=0$),
$u = u(x,t)$ = displacement of position $x$ at time $t$ relative to fixed laboratory coordinates,
$E$ = modulus of elasticity of projectile,
$\rho$ = density of projectile,
$A$ = cross-sectional area of projectile (as rod),
$F(t)$ = physically admissible force on leading face of projectile (as rod) at time $t$,
$V_0$ = initial velocity of projectile.

FIG. 6. Representative shadowgram of bullet during impact.

The function $u(x,t)$, which is the solution for (1) with the boundary conditions in (1a) and (1b), is given by the sum of the residues of the contour integral, see Fig. 9,

$$u(x,t) = \frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty} e^{\lambda t}\left[\frac{V_0}{\lambda^2} + \frac{c}{AE}\cdot\bar{F}(\lambda)\cdot\left(\cosh\frac{\lambda x}{c}\Big/\lambda\sinh\frac{\lambda l}{c}\right)\right]d\lambda, \quad (2)$$

prescribed by the inversion of a Laplace transform, where $\lambda$ is a complex variable, and

$$\bar{F}(\lambda) = \int_0^\infty e^{-\lambda t}F(t)dt \quad (3)$$

defines the Laplace transform of $F(t)$ assumed to exist for physically admissible forces in armor penetration.

The contour of Fig. 9 may have to be indented if $\bar{F}(\lambda)$ possesses any branch points. However, as long as the solution retains the property of uniform convergence, the sequel is unaffected.

The displacement, velocity, and deceleration of the projectile base are obtained at once by setting $x=0$ in (2) and these quantities are given respectively by,

$$u(0,t) = \frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty} e^{\lambda t}\left[\frac{V_0}{\lambda^2} + \frac{c}{AE}\cdot\left(\bar{F}(\lambda)\Big/\lambda\sinh\frac{\lambda l}{c}\right)\right]d\lambda, \quad (4)$$

$$\left.\frac{\partial u(x,t)}{\partial t}\right|_{x=0} = \frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty} e^{\lambda t}\left[\frac{V_0}{\lambda} + \frac{c}{AE}\cdot\left(\bar{F}(\lambda)\Big/\sinh\frac{\lambda l}{c}\right)\right]d\lambda, \quad (5)$$

and

$$\left.\frac{\partial^2 u(x,t)}{\partial t^2}\right|_{x=0} = \frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty} \frac{c}{AE}\cdot\left(\lambda\cdot e^{\lambda t}\cdot\bar{F}(\lambda)\Big/\sinh\frac{\lambda l}{c}\right)d\lambda. \quad (6)$$

The displacement, velocity, and deceleration of the center of mass are obtained from that of

the base by the following theorems proved in the Appendix:

*Theorem I:* The displacement, velocity, and deceleration of the center of mass at any instant $t'$ are respectively equal to the average values of the displacement, velocity, and deceleration of the projectile base over the time interval $(t'-l/c, t'+l/c)$ for all $-\infty < t' < \infty$.

*Theorem II:* The velocity of the center of mass at any instant $t'$ is equal to the difference between the displacements of the base at the end points of the time interval $(t'-l/c, t'+l/c)$ divided by the magnitude of the interval (the fundamental period of longitudinal vibration, $2l/c$).

*Theorem III:* The deceleration of the center of mass at any instant $t'$ is equal to the difference between the velocities of the base at the end points of the time interval $(t'-l/c, t'+l/c)$ divided by the magnitude of the interval, $2l/c$.

*Theorem IV:* The displacement, velocity, and deceleration of the center of mass at any instant are respectively equal to the average of the values of displacement, velocity, and deceleration distributed throughout the projectile at that instant.

Of these theorems it is seen that the fourth differs from the preceding three in that it involves an average over a position variable rather than time. It is the application of Theorem III to the primary slope (velocity) data, however, which constitutes the simplification brought about by the theory in the analysis of the shadowgrams. An earlier procedure of smoothing out the secondary maxima in the derived deceleration curves by a least square fitting is reduced to a simple differencing procedure which, to within the validity of the original assumption, is analytic. The striking feature of the arithmetic program implied in Theorem III is that it permits the evaluation of the forces back to the beginning of impact so that the forces which arise from impact at high speed may be compared with those from static punching.

The question has arisen during this investigation whether the force curves should be derived from comparator measurements of displacement rather than of slope (velocity). At present it is believed that greater precision in estimating the edge of the shadow on the film is necessary to settle this point. Attention is being directed now toward increasing film speed and reducing slit width to reduce the region of transition from the exposed to shadowed portions of the film strip to within the limitations imposed by grain size. Additional precision in the displacement data is

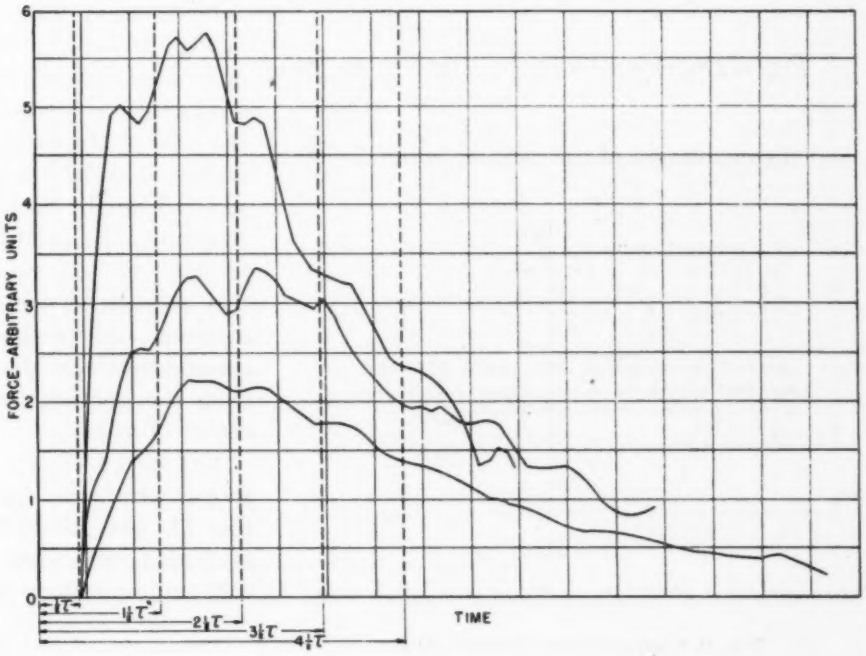



FIG. 7. Uncorrected force curves showing secondary maxima.

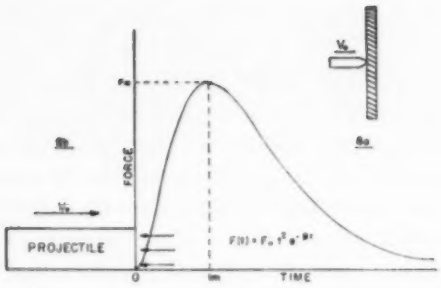



FIG. 8. Analytic representation of a projectile piercing armor.

desired since such data must be differentiated twice to obtain decelerations.

The formal derivation of the theorems given above has not required a specific formulation of the force function, $F(t)$, representing the forces which the projectile encounters during the impact. It has proved interesting to reverse the process to show the motion of the base of the projectile from an assumed force on the nose. Empirically, it can be shown that force curves (derived from the observed decelerations of the base such as are given in Fig. 7) are well depicted by the functional form

$$F(t)=F_0\cdot t^2\cdot e^{-\beta t}. \tag{7}$$

One then substitutes into (6) the transform

$$\bar{F}(\lambda)=2F_0/(\lambda+\beta)^3=F_0\int_0^\infty t^2\cdot e^{-(\lambda+\beta)t}dt, \tag{8}$$

and the evaluation of the contour integral yields

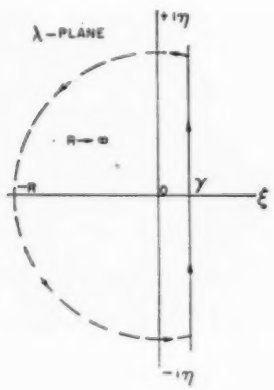



FIG. 9. Contour in the complex plane.

the series solution

$$\left.\frac{\partial^2 u(x,t)}{\partial t^2}\right|_{x=0}=F_0\cdot\frac{c}{AE}\cdot\frac{d^2}{d\lambda^2}\left.\left(\frac{\lambda e^{\lambda t}}{\sinh\frac{\lambda l}{c}}\right)\right|_{\lambda=-\beta} \tag{9}$$

$$-F_0\cdot\frac{4c}{AE}\cdot\sum_{n=1}^{\infty}\left[\frac{(-1)^n(n\pi c/l)^2}{n\pi\{\beta^2+(n\pi c/l)^2\}^3}\times\left(P_n\cdot\sin\frac{n\pi ct}{l}+Q_n\cdot\cos\frac{n\pi ct}{l}\right)\right],$$

where

$$P_n=\beta^3-3\beta\left(\frac{n\pi c}{l}\right)^2,\quad Q_n=\left(\frac{n\pi c}{l}\right)^3-3\beta^2\left(\frac{n\pi c}{l}\right),$$

and $t\geqslant 0$. It can be shown formally that the series given by (9) is representable by use of the Heaviside wave operators in the form

$$\left.\frac{\partial^2 u(x,t)}{\partial t^2}\right|_{x=0}=\frac{2c}{AE}\left[H(t-l/c)\cdot\frac{dF(t-l/c)}{dt}+H(t-3l/c)\cdot\frac{dF(t-3l/c)}{dt}+\cdots\right], \tag{10}$$

where

$$H(t-T)\equiv\begin{cases}0, & t<T\\ 1, & t>T\end{cases},\ \text{is the step operator.}$$

Normalized deceleration curves are presented in Fig. 10 showing the deceleration of the center of mass and of the trailing base for each of two steel projectiles when they encounter a force on the nose of the form $F_0\cdot t^2\cdot e^{-\beta t}$. The deceleration is plotted as the fractional value of the maximum suffered by the center of mass. The argument $\tau$ represents the fundamental period after impact for the one-inch projectile. Most striking from this calculation is the result that secondary maxima and minima under quite reasonable assumptions may deviate as much as 20 percent from the deceleration values attributed to the center of mass.

The effect of the shape of the projectile is shown with some representative force curves in Fig. 11. The curves result from the penetration of $\frac{1}{4}$-inch mild steel and STS[1] plates by two 0.27-calibre bullets of slightly different ogive.

[1] Special treated steel, 250 Brinell.

As should be expected, the force rises more slowly and comes to a narrower maximum for the sharp nosed than for the blunt nosed projectile. While the energy losses for the two are almost equal in just penetrating mild steel, the loss is greater for the blunter bullet in penetrating STS.

The application of the optical chronograph has recently been extended[2] to evaluate the effect of inertia and strain-rate at ballistic speeds by G. R. Irwin. Because the mathematical solution for the entry of an ogival-nosed bullet into armor presents exceedingly difficult geometrical features associated with the plastic flow of the deforming armor plate, Irwin employed conically nosed bullets. It was shown in this study that the experimental precision is good enough to permit estimation of the magnitude of velocity effects in rapid deformation of steel. The paper by Irwin should be consulted for further details.


## ACKNOWLEDGMENT

The optical chronograph was successfully developed at the Naval Research Laboratory prior to the writer's association with the problem, by G. R. Irwin, Frank Louckes, and G. D. Kinzer. Valuable assistance in the taking and analyzing of shadowgrams was rendered by Jean R. Streeter and Jack Hodges. Reviewing comments of W. H. Sanders and Jean Streeter were very helpful.


## APPENDIX

The proof for the four theorems given in the text are based on the solution given in Eq (2):

$$u(x,t)=\frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty} e^{\lambda t}\left[\frac{V_0}{\lambda^2}+\frac{c}{AE}\cdot\bar{F}(\lambda)\cdot\frac{\cosh(\lambda x/c)}{\lambda\sinh(\lambda l/c)}\right]d\lambda,$$

which gives the displacement $u$ with respect to fixed laboratory coordinates of the cross section located at $0\leq x\leq l$ in the moving projectile for the time $t\geq 0$

For Theorem I, we need give the proof only for the displacement part of the theorem since the proofs for velocity and acceleration involve trivial differences From the expression above and from (4), we find the average:

$$\begin{aligned}
&\frac{c}{2l}\int_{t'-l/c}^{t'+l/c} u(0,t)\cdot dt\\
&=\frac{c}{2l}\cdot\frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty}\left\{\int_{t'-l/c}^{t'+l/c} e^{\lambda t}dt\right\}\left\{\frac{V_0}{\lambda^2}+\frac{c}{AE}\cdot\frac{\bar{F}(\lambda)}{\lambda\sinh(\lambda l/c)}\right\}d\lambda\\
&=\frac{c}{l}\cdot\frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty}\left[e^{\lambda t'}V_0\frac{\sinh(\lambda l/c)}{\lambda^3}+\frac{c}{AE}\cdot e^{\lambda t'}\cdot\frac{\bar{F}(\lambda)}{\lambda^2}\right]d\lambda\\
&=V_0t'+\frac{1}{M}\cdot\frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty}\left\{\int_{-\infty}^{t'}\left(\int_{-\infty}^{t'} e^{\lambda t}dt\right)dt'\right\}\bar{F}(\lambda)d\lambda\\
&=V_0t'+\frac{1}{M}\int_{-\infty}^{t'}\left(\int_{-\infty}^{t'}\left[\frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty} e^{\lambda t}\bar{F}(\lambda)d\lambda\right]dt\right)dt'\\
&=V_0t'+\frac{1}{M}\int_{-\infty}^{t'}\left(\int_{-\infty}^{t'} F(t)dt\right)dt',
\end{aligned}$$

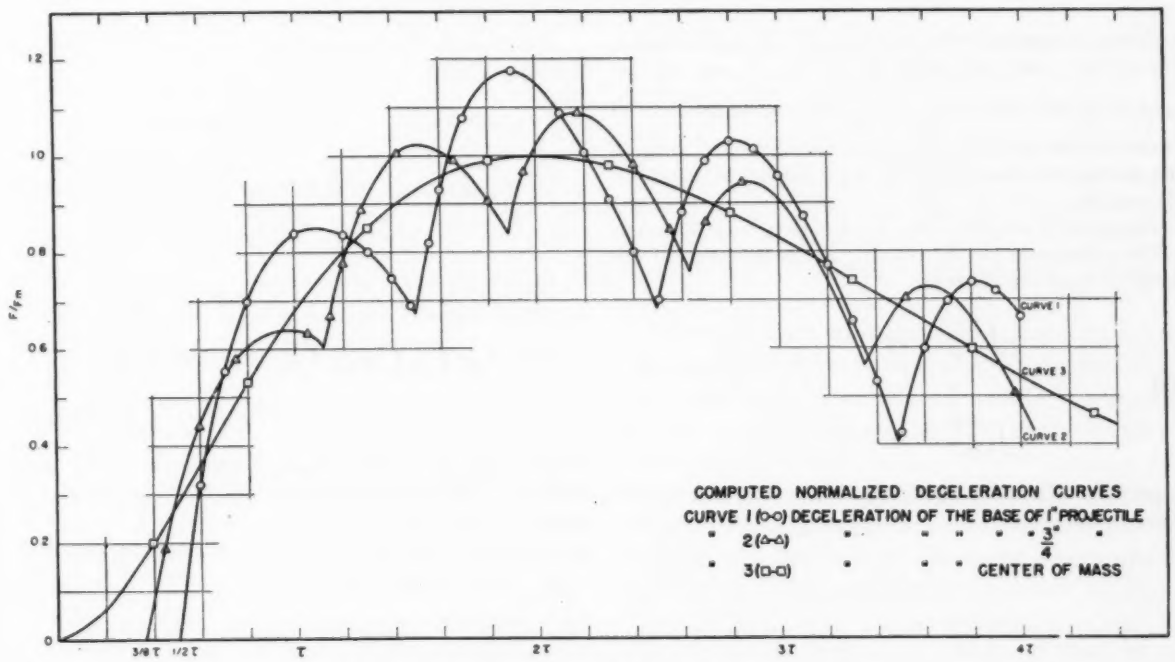



FIG. 10. Computed normalized deceleration curves.

[2] G. R. Irwin, "Penetration resistance measurements at ballistic speeds," Proceedings of the Sixth International Congress for Applied Mechanics, September, 1946.

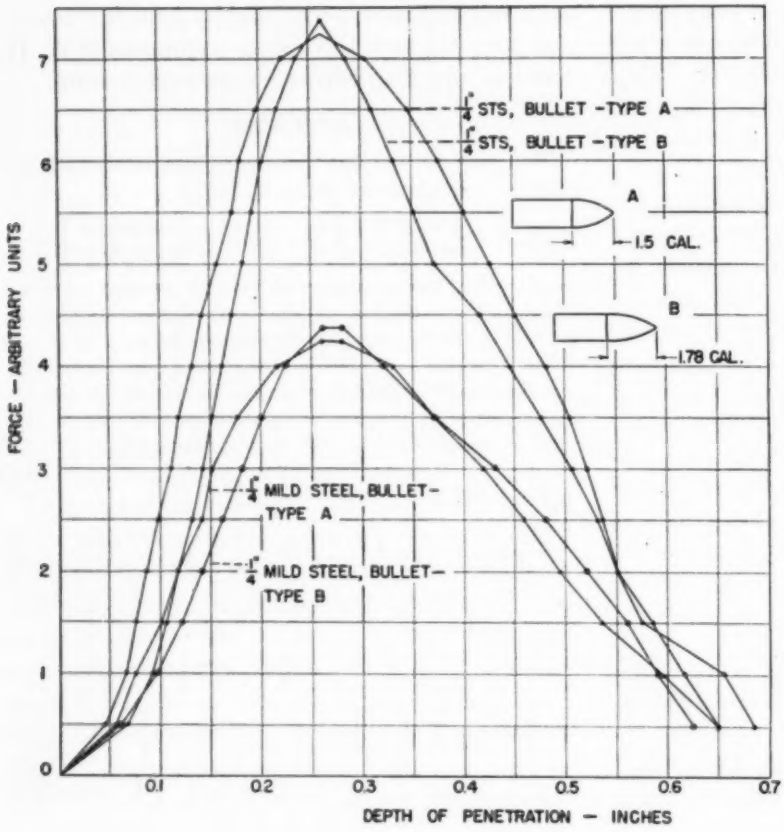



FIG. 11. Effect of projectile shaep on force curves.

and this is precisely the displacement of a point-mass, $M = AEl/c^2$, which has an initial velocity $V_0$ and is acted upon by a force $F(t) = \begin{cases} 0, & t<0 \\ F(t), & t \geq 0 \end{cases}$ up to the instant $t'$. The proofs for the velocity and deceleration similarly involve the taking of averages with the expressions given in (5) and (6).

Theorems II and III follow immediately from Theorem I. The expressions for the velocity and deceleration of the center of mass are given by

$$\frac{c}{2l}\int_{t'-l/c}^{t'+l/c} \partial u(x,t)/\partial t\Big|_{x=0} dt$$

and

$$\frac{c}{2l}\int_{t'+l/c}^{t'-l/c} \partial^2 u(x,t)/\partial t^2\Big|_{x=0} dt$$

respectively. These may be integrated at once to give the desired expressions.

For Theorem IV the proof need be given only for the case of displacement. The average is

$$\begin{aligned}
&\frac{1}{l}\int_0^l u(x,t)\cdot dx \\
&= \frac{1}{l}\cdot\frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty}\left\{\int_0^l e^{\lambda t}\left[\frac{V_0}{\lambda^2}+\frac{c}{AE}\cdot\frac{\bar{F}(\lambda)\cdot\cosh(\lambda x/c)}{\lambda\sinh(\lambda l/c)}\right]dx\right\}d\lambda \\
&= \frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty} e^{\lambda t}\left[\frac{V_0}{\lambda^2}+\frac{1}{M}\cdot\frac{\bar{F}(\lambda)}{\lambda^2}\right]d\lambda \\
&= V_0 t+\frac{1}{M}\cdot\frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty}\bar{F}(\lambda)\left[\int_{-\infty}^{t} dt\int_{-\infty}^{t} e^{\lambda t}dt\right]d\lambda \\
&= V_0 t+\frac{1}{M}\int_{-\infty}^{t}\left(\int_{-\infty}^{t}\left[\frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty} e^{\lambda t}\bar{F}(\lambda)d\lambda\right]dt\right)dt \\
&= V_0 t+\frac{1}{M}\int_{-\infty}^{t}\left(\int_{-\infty}^{t} F(t)dt\right)dt,
\end{aligned}$$

which is identical with the result in Theorem I for the upper limit $t = t'$. The proofs for velocity and acceleration follow at once.

# Probability Criterion for the Design of Servomechanisms

JOHN R. RAGAZZINI AND LOTFI A. ZADEH
*Department of Electrical Engineering, Columbia University, New York, New York*
(Received March 31, 1948)

Of the criteria that are used in the design of servomechanisms, the so-called "minimum mean-square" error criterion is used quite widely in cases where the undesirability of the error increases with its magnitude. There is, however, an important class of applications, particularly in the field of ballistics, where a more appropriate criterion can be defined as the probability that the error be less than some prescribed tolerance. The problem of maximization of this probability is shown to reduce to a number of implicit equations whose solution in practice must be carried out by cut and trial. It is further shown that under certain conditions the maximization of probability yields the same values for the design constants as the minimization of mean-square error. These conditions are discussed for the general case as well as for certain specific cases which are important in practice.

## I. INTRODUCTION

THE design of servomechanisms is usually based on some generally accepted criterion of performance. A relatively simple and yet useful one which was introduced some years ago[1] is based on minimization of the mean-square error. This criterion is applicable in situations where the undesirability of the error increases with its magnitude. While this is the case in most common applications, there is an important class of problems in ballistics and certain industrial applications where it is desirable that the magnitude of the error, $\epsilon$, be less than a certain critical value, $L$. Thus, in these problems, all errors larger than $L$ are equally bad, while those smaller than $L$ are equally acceptable. It is apparent that for such cases the mean-square error criterion does not reflect correctly the requirements of the system.

As a measure of performance for this class of problems, the authors have been using a figure of merit, $p$, which is defined by the following functional

$$p = \text{Prob.}\{|\epsilon| < L\}. \tag{1}$$

The difficulty arising in the application of (1) is that the complexity of the functional usually prohibits purely analytical approaches such as are possible in the case of the mean-square error criterion.[2–4] However, it is shown in this paper that in many problems where the *a priori* requirement of the design is that $p$ be maximum, the relative values of systematic error, random error, and the tolerance $L$ are such that practically the same design is obtained as through application of the mean-square error criterion. In the following sections a criterion for the equivalence of the two design procedures is obtained from comparison of the analytic conditions for maximization of $p$ and minimization of the mean-square error.

## II. MAXIMIZATION OF THE PROBABILITY FIGURE $p$

A typical servomechanism which shall be considered in these discussions is shown in Fig. 1. The principal function of the device is to close out the error, $\epsilon(t)$. The performance of the system is judged by its success (or failure) to bring the error to zero at one or more specified instants of time or on a time average basis.

The structure and component values of the servomechanism are specified completely except for a number of adjustable parameters $\alpha_1, \alpha_2, \cdots \alpha_n$. Referring to Fig. 2, the error $\epsilon(t)$ results from a non-random input function $\theta_i(t)$ which is zero for $t<0$. It is assumed that as a result of random disturbances arising both inside and outside the system, the error $\epsilon$ comprises a systematic (non-random) component $\epsilon_s$, and a

[1] A. Kharkevitch, "On the application of the criterion of mean-square error to the evaluation of distortion in linear systems," J. Tech. Phys. 7, 515, 1065 (1937).

[2] N. Wiener, *The Extrapolation, Interpolation and Smoothing of Stationary Time Series*, NDRC Report 6037, Feb. 1, 1942.

[3] N. Levinson, *The Wiener RMS Error Criterion in Filter Design and Prediction*, J. Math. and Phys. 25, 261 (1947).

[4] Radiation Laboratory Series, *Theory of Servomechanisms* (McGraw-Hill Book Company, Inc., New York, 1947), pp. 308 *et seq.*

stationary random (noise) component $\epsilon_N$ so that

$$\epsilon = \epsilon_s + \epsilon_N. \tag{2}$$

$\epsilon_N$ is distributed normally with zero mean and a variance $\sigma^2$ which in view of the stationariness of $\epsilon_N$ is equal to its mean-square value. These assumptions are found to be generally true in practice.

The object of the design is to assign values to $\alpha_i$ which maximize the probability that at a prescribed instant $t=t_0$, the magnitude of $\epsilon$ would be less than a given tolerance $L$, as shown in Fig. 2. In short, the requirement is that

$$\frac{\partial p_0}{\partial \alpha_i} = 0; \quad i = 1, 2, 3, \cdots, n \tag{3}$$

where

$$p_0 = \text{Prob}\{|\epsilon|_{t=t_0} < L\}. \tag{4}$$

These conditions are subject, of course, to the physical realizability of the system.

As stated, the problem requires that $\alpha_1, \alpha_2 \cdots \alpha_n$ be determined on the basis of maximization of $p_0$. With a normally distributed noise variable the probability $p_0$ may be written as

$$p_0 = \frac{1}{(2\pi)^{\frac{1}{2}}\sigma} \int_{\epsilon_s^0 - L}^{\epsilon_s^0 + L} \exp\left(-\frac{1}{2}\frac{u^2}{\sigma^2}\right) du, \tag{5}$$

where $u$ is the variable of integration and $\epsilon_s^0$ denotes the value of the systematic error at time $t=t_0$. For any particular type of input $\theta_i$ which is selected as a test function, both $\epsilon_s^0$ and $\sigma$ can be expressed as functions of $\alpha_1, \alpha_2 \cdots \alpha_n$. Thus Eq. (5) may be regarded as expressing $p_0$ as an explicit function of $\alpha_1, \alpha_2 \cdots \alpha_n$. To find the equations satisfied by the maximizing values of $\alpha_i$, it is sufficient to differentiate the right-hand member of (5) partially with respect to the $\alpha$'s and to equate the derivatives to zero. The results of this differentiation are

$$(2\pi)^{\frac{1}{2}} \frac{\partial p_0}{\partial \alpha_i} = -\frac{1}{\sigma}\frac{\partial \epsilon_s^0}{\partial \alpha_i}\left\{\exp\left[-\frac{(\epsilon_s^0 + L)^2}{2\sigma^2}\right] - \exp\left[-\frac{(\epsilon_s^0 - L)^2}{2\sigma^2}\right]\right\} + \frac{\partial \sigma}{\partial \alpha_i}\frac{1}{\sigma^2} \times \int_{\epsilon_s^0 - L}^{\epsilon_s^0 + L} (u^2/\sigma^2 - 1) \times \exp\left[-\frac{u^2}{2\sigma^2}\right] du. \tag{6}$$

The coefficient of $\partial\sigma/\partial\alpha_i$ in (6) is integrable and can easily be shown to be equal to

$$\frac{1}{\sigma^2}\int_{\epsilon_s^0 - L}^{\epsilon_s^0 + L} (u^2/\sigma^2 - 1) \exp\left[-\frac{u^2}{2\sigma^2}\right] du = \frac{(\epsilon_s^0 - L)}{\sigma^2} \exp\left[-\frac{(\epsilon_s^0 - L)^2}{2\sigma^2}\right] - \frac{(\epsilon_s^0 + L)}{\sigma^2} \exp\left[-\frac{(\epsilon_s^0 + L)^2}{2\sigma^2}\right]. \tag{7}$$

Thus, finally, the equations satisfied by the maximizing values of $\alpha_1, \alpha_2 \cdots \alpha_n$ take the form

$$\frac{\partial \epsilon_s^0}{\partial \alpha_i}\left\{\exp\left[-\frac{(\epsilon_s^0 + L)^2}{2\sigma^2}\right] - \exp\left[-\frac{(\epsilon_s^0 - L)}{2\sigma^2}\right]\right\} - \frac{\partial \sigma}{\partial \alpha_i}\left\{\frac{(\epsilon_s^0 + L)}{\sigma} \exp\left[-\frac{(\epsilon_s^0 + L)^2}{2\sigma^2}\right] - \frac{(\epsilon_s^0 - L)}{\sigma} \exp\left[-\frac{(\epsilon_s^0 - L)^2}{2\sigma^2}\right]\right\} = 0. \tag{8}$$

The complexity of these equations is evident, and their solution can be arrived at only by cut and trial. It will be shown in the next section that under certain circumstances the minimum mean-square error criterion results in the same values for the control coefficients $\alpha_i$, as would be obtained from solution of (8).

## III. COMPARISON OF MAXIMUM PROBABILITY AND MINIMUM MEAN-SQUARE ERROR CRITERIA

In order to make comparisons between the two criteria, an expression similar to (8) will be obtained for the case of the minimum mean-square error criterion. First it should be noted that since the signal is a non-repetitive transient and the error must be minimized at $t=t_0$, the conventional minimum mean-square error criterion cannot be applied without some modifications. It is necessary to regard $E^2$ not as the average of $\epsilon^2$ over a long period of time, but rather as an average of $\epsilon^2$ over the ensemble. With this understanding, the mean-square value of the

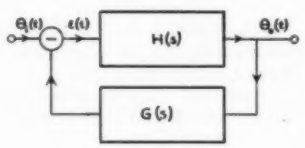



FIG. 1. Schematic diagram of a simple servomechanism.

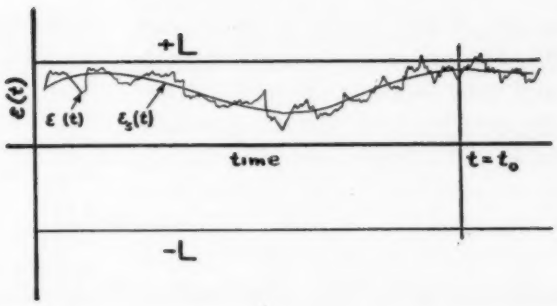



FIG. 2. Error function of a servomechanism.

error takes the form

$$E^2=(\epsilon_s^0)^2+\sigma^2, \tag{9}$$

and the minimizing values of $\alpha_i$ satisfy the equations

$$\frac{\partial E^2}{\partial \alpha_i}=0, \tag{10}$$

that is

$$\frac{\partial \epsilon_s^0}{\partial \alpha_i}\epsilon_s^0+\frac{\partial \sigma}{\partial \alpha_i}\sigma=0; \quad i=1, 2, \cdots, n. \tag{11}$$

These equations are the analogs of Eqs. (8) which were obtained on the basis of maximization of $p_0$. It is the relative simplicity of Eqs. (11), as contrasted to the complexity of Eqs. (8), that makes it possible in many instances to obtain solutions to (11) by purely analytic means.

In order that the solutions of Eqs. (11) be equal to the solutions of (8), it is necessary that the coefficients of $\partial\epsilon_s^0/\partial\alpha_i$ and $\partial\sigma/\partial\alpha_i$ be proportional to each other. That is

$$\frac{\exp\left[-\frac{(\epsilon_s^0+L)^2}{2\sigma^2}\right]-\exp\left[-\frac{(\epsilon_s^0-L)^2}{2\sigma^2}\right]}{\epsilon_s^0}=\frac{\frac{(\epsilon_s^0+L)}{\sigma}\exp\left[-\frac{(\epsilon_s^0+L)^2}{2\sigma^2}\right]-\frac{(\epsilon_s^0-L)}{\sigma}\exp\left[-\frac{(\epsilon_s^0-L)^2}{2\sigma^2}\right]}{\sigma}. \tag{12}$$

After some manipulation, the above relation reduces to the following:

$$\exp\left[-\frac{2\epsilon_s^0 L}{\sigma^2}\right]=\frac{1-\frac{\epsilon_s^0 L}{\sigma^2}+\frac{(\epsilon_s^0)^2}{\sigma^2}}{1+\frac{\epsilon_s^0 L}{\sigma^2}+\frac{(\epsilon_s^0)^2}{\sigma^2}}. \tag{13}$$

Within the limitations of the assumptions made in the statement of the problem, Eq. (13) represents the necessary and sufficient condition for the equivalence of the minimum mean-square error and maximum probability criteria.

## IV. DISCUSSION OF RESULTS

An examination of (13) shows that there are only two situations of practical interest where this equation is satisfied approximately over a wide range of values of $L$, $\epsilon_s^0$, and $\sigma$. These conditions are

$$\left.\begin{aligned}|\epsilon_s^0| &\ll \sigma \\ (L\epsilon_s^0)^{\frac{1}{2}} &\ll \sigma\end{aligned}\right\} \tag{14}$$

or

$$\left.\begin{aligned}|\epsilon_s^0| &\gg \sigma \\ |\epsilon_s^0| &\cong L.\end{aligned}\right\} \tag{15}$$

Stated in words, the mean-square error and maximum probability criteria are approximately equivalent when either:

(a) The systematic error as well as the geometric mean of the tolerance $L$ and the systematic error at the time $t=t_0$ are small in comparison with the r.m.s. value of the random error.

(b) The systematic error is large in comparison with the random error, and the tolerance $L$ is approximately equal to the systematic error.

## V. CONCLUSION

In the design of that class of servomechanism in which it is desired to close out the error within a prescribed tolerance at a specific instant such as at the termination of a given operation, it is necessary to maximize the probability that the error at that instant be less than the given tolerance. This can be done by adjustment of the

dynamic constants of the system in accordance with Eqs. (8). The process is generally quite complicated and must be carried out by cut and trial. Under the general conditions expressed by (13) and in particular by (14) and (15), the resultant design is the same as that which would be obtained by the use of the minimum mean-square error criterion. In view of the relative simplicity of the latter, considerable economy in computational labor can thus be achieved.

## Measurements of the Mechanical Properties of Polymer Solutions by Electromagnetic Transducers*

THOR L. SMITH,** JOHN D. FERRY, AND FREDERIC W. SCHREMP
*Department of Chemistry, University of Wisconsin, Madison, Wisconsin*

(Received June 23, 1948)

A concentrated polymer solution is sheared by a rod oscillating axially with a very small amplitude in a closed tube. The rod is driven by a coil in a magnetic field; from electrical measurements on the coil, the mechanical resistance and reactance of the system are calculated by transducer relationships. Equations are given for obtaining the dynamic viscosity, $\eta'$, and rigidity, $G'$, of the solution. In the case of some polymer solutions, the audiofrequency range covered by the present apparatus falls in the dispersion zone where $\eta'$ and $G'$ change rapidly with frequency. In other cases the frequency range covered appears to be at the upper end of the dispersion zone; $\eta'$ is far smaller than the viscosity in steady flow, and $G'$ changes very little with frequency. The advantages and limitations of the method are discussed.

### INTRODUCTION

AMONG the diverse phenomena exhibited in the mechanical properties of concentrated polymer solutions, their behavior in small oscillating deformations is of particular interest.[1] At a single frequency of oscillation, the mechanical behavior of such a solution can be simply described by two quantities, the dynamic rigidity and dynamic viscosity. In general, both these quantities are found to be frequency-dependent, and in many cases their dispersion is spread over a wide range of frequencies which cannot be covered by a single experimental method. A variety of methods is required for an adequate study.

Measurements of transverse wave propagation[2] have proved useful for obtaining the rigidity, and under favorable conditions also the viscosity, in the audiofrequency range, provided the mechanical loss factor of the solution is not so great that the waves are too severely damped. The new method described in this paper provides a better determination of viscosity and is applicable for both viscosity and rigidity under some conditions where wave measurements fail owing to extreme damping.

If non-linear effects in either rigidity or viscosity can safely be ignored, as they usually are in treatments of viscoelastic systems,[1,3] both shear and rate of shear must be kept small throughout the cycle of deformation. Fulfilment of this requirement is facilitated by our present apparatus, in which neither shear strain nor stress is measured; the only measurements are of electrical resistance and capacitance, and the viscosity and rigidity are calculated from transducer relationships. In this respect our apparatus differs from that of Philippoff,[4] who employed a somewhat similar device for shearing a solution but calculated the viscosity from measurements of the amplitude of oscillation at mechanical resonance.

* Presented at the Division of High Polymer Physics of the American Physical Society, Chicago, December 30, 1947.

** Present address, Hercules Powder Company, Wilmington, Delaware.

[1] J. D. Ferry, W. M. Sawyer, and J. N. Ashworth, J. Polymer Sci. 2, 593 (1947).

[2] J. D. Ferry, J. Am. Chem. Soc. 64, 1323 (1942); J. N. Ashworth, W. M. Sawyer, and J. D. Ferry, unpublished experiments.

[3] T. Alfrey and P. Doty, J. App. Phys. 16, 700 (1945).

[4] W. Philippoff, Physik. Zeits. 35, 884, 900 (1934).

## DESCRIPTION OF APPARATUS

An aluminum rod, 3.18 mm in diameter, is driven axially in a tube, shearing the solution to be studied across an annular gap of about 1 mm. The maximum excursion of the rod is always less than 0.1 mm, and often less than 0.01 mm. The top of the rod is screwed and cemented into a cylinder of Lucite, which is machined to leave a thin shell for a coil form. Several hundred turns of No. 34 enameled copper wire are wound on the shell, which is suspended in the air gap of a 10-inch permanent magnet loudspeaker.

Two models of the apparatus are shown in Figs. 1 and 2. In Model 1, suitable for a frequency range of about 30 to 200 cycles/sec., the rod and coil are supported by a brass disk attached to four strands of No. 005 music wire, 10 cm in length; the coil is centered with equal tension in the four wires by hanging equal weights over pulleys, as shown, before the clamps are tightened. In Model 2, suitable for a frequency range of about 200 to 500 cycles/sec., the rod and coil are supported by a brass disk silver-soldered to four pieces of $\frac{1}{16}$-in. drill rod. The outer ends of the rods are soldered in holes drilled through a heavy brass ring. The effective length of the rods in flexure is 2.2 cm. It is essential that the moving system be very rigid, although light, and that there be no friction where the coil moves in the magnetic field.

In each model, the tube is of glass closed by a brass plug, sealed in with litharge and glycerine cement, and surrounded by a water jacket. The tube and jacket can be lowered for filling and emptying; when in use, they are clamped with the end of the rod about 1 mm above the plug, the depth of immersion being varied by changing the volume of solution in the tube. The clamping is self-aligning, and the whole mounting must be extremely rigid; earlier designs with less massive construction showed anomalous frequency-dependent phenomena which were traced to slight compliances in the mounting. The vertical support rods are screwed tightly into a heavy base which rests on sponge rubber.

The coil is made one arm of an impedance bridge (Fig. 3). The driving current is usually 0.01–0.02 amp. The oscillator is a Hewlett-Packard Model 200 CR;[5] the amplifier, sometimes omitted, is a conventional public address system type; $R_2$ is a non-inductive resistance, 50 or 75 ohms; $R_3$ is two General Radio[6] type 500-H resistances, each 1000 ohms, connected in parallel;

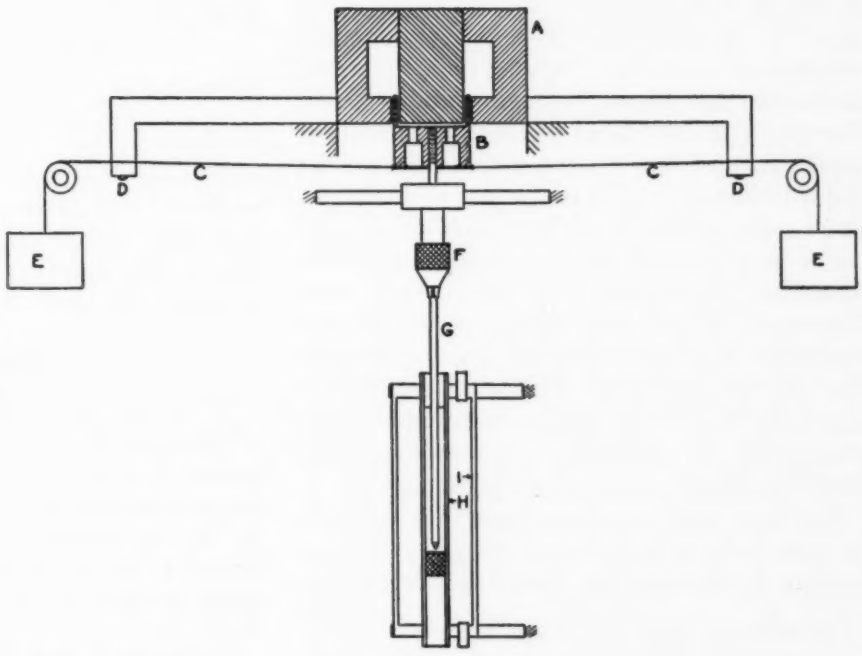



FIG. 1. Schematic diagram of Model 1 transducer. A, loudspeaker; B, Lucite coil form; C, suspension wires; D, clamps; E, weights for adjusting tension; F, chuck; G, aluminum rod; H, tube; I, water jacket. Supports of speaker and water jacket are not shown.

[5] Hewlett-Packard Company, Palo Alto, California.
[6] General Radio Company, Cambridge, Massachusetts.

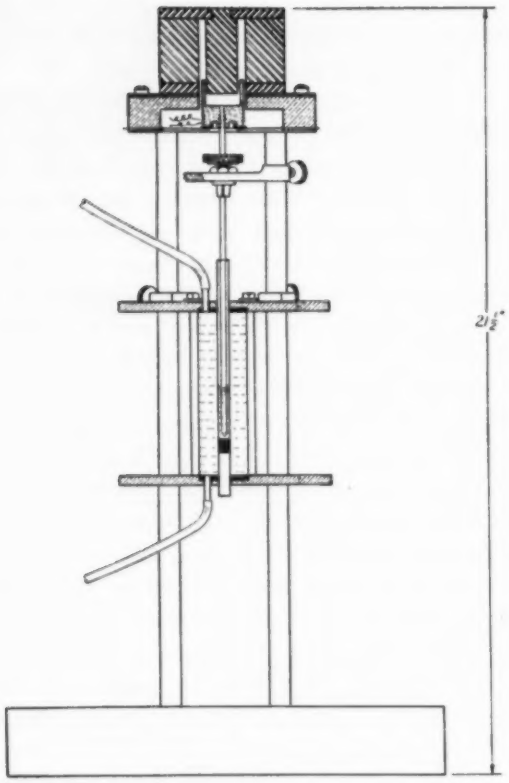


FIG. 2. Cross-section drawing of transducer Model 2.

$R_4$ is a series of decade resistances, General Radio type 510, with additional plug-in resistances and a 1-ohm slide wire, totalling 3000 ohms; $C$ is a set of decade capacitances, General Radio type 380, together with a 0.003 $\mu$f variable air condenser, totalling 2 $\mu$f. When the reactance of the moving coil is inductive, $C$ is placed in parallel with $R_4$; when capacitative, in parallel with $R_3$. The detector is a General Radio Type 760A Sound Analyzer, preceded by a Kay-Lab Type 102A Decade Amplifier.[7] The resistance, $R$, and the reactance, $X$, of the moving coil are calculated by the usual methods of a.c. bridges, as described in detail elsewhere.[8] The current through the coil is determined by measuring the potential drop across $R_2$ with a vacuum tube voltmeter.

For the transducer calculations, it is necessary to have either the impedance of the coil when motion is prevented by clamping,[9] or the impedance with various masses attached at constant frequency.[10] Clamping is achieved by a chuck (Fig. 1) or a collet (Fig. 2). Variation of mass, which is useful only with Model 2 (because added mass in Model 1 would change the tension in the wires and shift the equilibrium position of the coil to a region of different magnetic field intensity), is achieved by attaching small weights to the upper part of the rod with set screws.

## TRANSDUCER CALCULATIONS

### Method 1

The resistance and reactance of the coil are measured while the rod is oscillated ($R$, $X$) and while it is rigidly clamped ($R_0$, $X_0$). The fundamental electromagnetic transducer equation[9] is $\mathbf{Z_M}=(Bl)^2\times10^{-9}/(\mathbf{Z}-\mathbf{Z_0})$, where $\mathbf{Z}$ and $\mathbf{Z_0}$ are the electrical impedances in motion and clamped, and $\mathbf{Z_M}$ is the mechanical impedance of the system, defined as $\mathbf{f}/\dot{x}$, $\mathbf{f}$ being the (complex) driving force and $\dot{x}$ the velocity. In the constant $(Bl)^2\times10^{-9}$, which has the practical units ohm-dyn-sec.-cm$^{-1}$, $B$ is the magnetic field strength in gauss, and $l$ is the length of wire in the coil. Separation into real and imaginary parts yields the equations used for calculation:

$$R_M=(Bl)^2(R-R_0)\times10^{-9}/[(R-R_0)^2+(X-X_0)^2], \quad (1)$$

$$X_M=-(Bl)^2(X-X_0)\times10^{-9}/[(R-R_0)^2+(X-X_0)^2], \quad (2)$$

where $\mathbf{Z_M}=R_M+iX_M$. This method is always employed with Model 1; the value of $R_0$ is about 10.6 ohms, and $R-R_0$ ordinarily ranges from 0.1 ohm to several ohms, measured with a precision of about 0.01 ohm.

### Method 2[11]

When the mass of the moving system is varied at constant frequency, the locus of $X$ plotted against $R$ is a circle,[10] and the coordinates of its left intersection with its horizontal diameter are $R_0$, $X_0$. For Model 2, $R_0$ is about 24.7 ohms, and $R$ and $X$ are ordinarily determined with a precision of about 0.02 ohms. The value of $R_M$ is

[7] Kalbfell Laboratories, San Diego, California.
[8] T. L. Smith, Ph.D. thesis, University of Wisconsin, 1948.
[9] W. P. Mason, *Electromechanical Transducers and Wave Filters* (D. Van Nostrand and Company, Inc., 1942).
[10] R. D. Fay, J. Acous. Soc. Am. 15, 32 (1943).
[11] We are indebted to Professor R. D. Fay, Massachusetts Institute of Technology, for calling our attention to this method.

most readily calculated by geometry as $(Bl)^2 \times 10^{-9}/2r$, where $r$ is the radius of the circle; $X_M$ is obtained from the coordinates of any single point on the circle, together with $R_0$ and $X_0$, using Eq. (2). It is unnecessary to know the magnitude of the added masses; they are simply selected to space the points suitably around the circle. The points move in a clockwise direction with increasing mass. It is impractical to make these plots at frequencies much above the mechanical resonance frequency corresponding to zero added mass, since here the points are limited to the lower half of the circle. Examples of this type of plot are shown in Fig. 4.

### Determination of the Constant $(Bl)^2 \times 10^{-9}$

When measurements are made with the rod moving freely in air, $X_M = \omega M - S_M{}^0/\omega$, where $\omega$ is the circular frequency, $M$ is the effective mass of the moving system, and $S_M{}^0$ is the elastance in dyne/cm of the mounting. Accordingly, if measurements are made at different frequencies and the quantity $-(X-X_0)/[(R-R_0)^2+(X-X_0)^2]\omega$ is plotted against $1/\omega^2$, the intercept is $M/(Bl)^2 \times 10^{-9}$, and the slope is $-S_M{}^0/(Bl)^2 \times 10^{-9}$. If there is any uncertainty about the effective mass of the rod and coil, $(Bl)^2 \times 10^{-9}$ can be determined from the change in intercept when known masses are attached. The values for this constant in ohm-dyn-sec.-cm$^{-1}$ are $1.06 \times 10^4$ in Model 1 and $1.49 \times 10^4$ in Model 2. The value of $S_M{}^0$ for Model 2 is approximately $8 \times 10^7$ dyn cm$^{-1}$; for Model 1 it is extremely small.

### CALCULATION OF VISCOSITY AND RIGIDITY

The dynamic viscosity $\eta'$ is defined as the component of shear stress in phase with the rate of strain divided by the rate of strain, and the dynamic rigidity $G'$ the component of stress in phase with the strain divided by the strain.[1] It follows that

$$\eta' = R_M/a, \tag{3}$$

and

$$G' = (\omega^2 M - \omega X_M - S_M{}^0)/a, \tag{4}$$

where $a$ is a constant with dimensions of length which depends on the geometrical details of the shearing.[1]

We originally employed a U-tube with an open end, following the example of Philippoff.[4] If the mechanical resistance of a rod moving axially in a U-tube is assumed to be given by the viscous drag of the liquid between cylinders plus an end effect independent of immersion, then[12] $R_M = 2\pi\eta' L/\ln(r_a/r_i) + c$, where $L$ is the depth of immersion, $r_a$ the radius of the tube, $r_i$ that of the rod, and $c$ the end effect; thus

$$\eta' = (dR_M/dL)\ln(r_a/r_i)/2\pi, \tag{5}$$

and it can be determined from the slope of the upper part of the plot of $R_M$ against $L$ (Fig. 5). In agreement with the experience of Philippoff, we found that the end effect was enormous; moreover, it was not represented by a clear-cut intercept of a linear plot; instead, the line curved down for small values of $L$ (Fig. 5), so it was difficult to select a truly linear segment of the

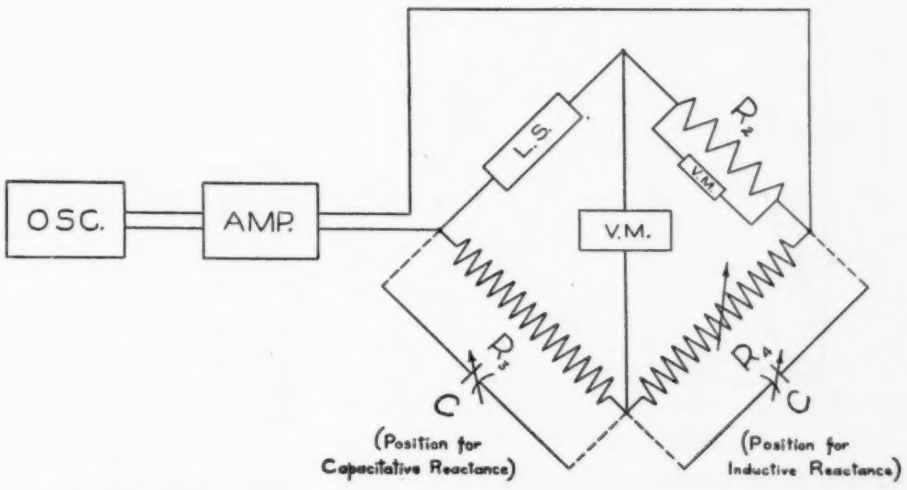



FIG. 3. Impedance bridge for transducer. L.S. is loudspeaker coil.

[12] A. Pochettino, Nuovo Cimento **8**, 77 (1914).

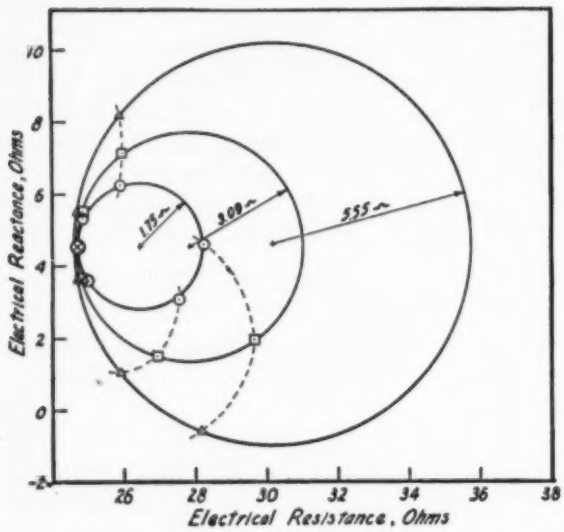



FIG. 4. Examples of circle plots for 20.2 percent polyvinyl acetate in cyclohexanone, 25°C, frequency 320 cycles. ○, $L=4.0$, $R_M=0.427\times10^4$; □, $L=2.2$, $R_M=0.241\times10^4$; △, $L=1.2$, $R_M=0.134\times10^4$.

upper part of the curve. These difficulties are avoided, and the sensitivity of the method is greatly increased, by using a closed tube as described above.

The explanation of the curvature in the case of the U-tube is qualitatively as follows. With deep immersion the liquid displaced by descent of the rod is largely forced around the bend, but with decreasing immersion some of it flows back up around the rod, resulting in a greater mechanical resistance per unit immersion. For very small immersions all the displaced liquid flows up, the effect being similar to that in a closed tube, and indeed the initial slopes for U- and closed tubes are approximately identical (Fig. 5).

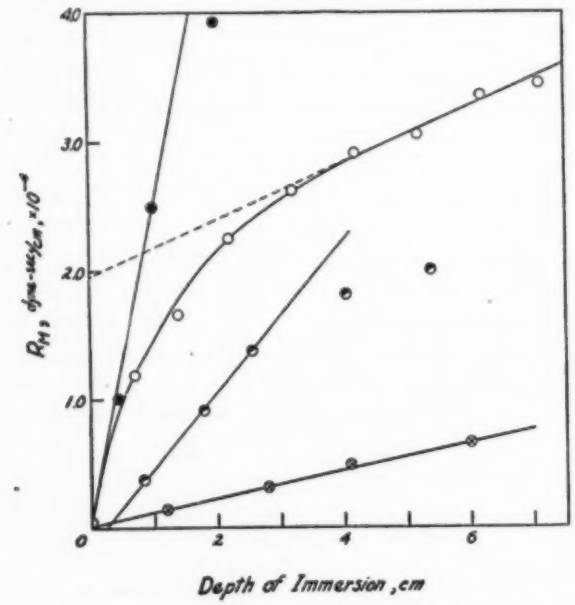



FIG. 5. $R_M$ plotted against $L$ for oils and honey. ○, U-tube, oil OB-1; ●, closed tube, oil OB-1; ◐, closed tube, honey; ⊗, closed tube, oil N-13.

For the closed tube, all the displaced liquid is of course forced back up. In this case, the liquid near the descending rod moves down, while farther out near the wall of the tube it moves up. The velocity $v$ and velocity gradient $\partial v/\partial r$ as functions of radial distance are given by

$$v=v_i\frac{(r_a^2+r_i^2)\ln(r/r_a)+(r_a^2-r^2)}{(r_a^2+r_i^2)\ln(r_a/r_i)-(r_a^2-r_i^2)}, \quad (6)$$

$$\frac{\partial v}{\partial r}=v_i\frac{r_a^2+r_i^2-2r^2}{r[(r_a^2+r_i^2)\ln(r_a/r_i)-(r_a^2-r_i^2)]}, \quad (7)$$

where the velocity of the rod is $-v_i$. These equations are derived in an Appendix. A plot of the functions is shown in Fig. 6. At $r=r_0$ the velocity is 0, and at $r=r_1$ the velocity gradient is 0. Neglecting any end effect, it follows (see Appendix) that

$$R_M=\frac{2\pi\eta' L(r_a^2+r_i^2)/(r_a^2-r_i^2)}{(r_a^2+r_i^2)\ln(r_a/r_i)/(r_a^2-r_i^2)-1}=a'\eta' L, \quad (8)$$

where $a'$ is now a dimensionless apparatus constant.

When $R_M$ is plotted against $L$, the line obtained usually strikes the origin; if a small departure is observed, $\eta'$ is calculated as $(1/a')dR_M/dL$. No particular significance is attached to the value of the intercept, which may reflect a small residual mechanical resistance ($R_M^0$) of the apparatus, as well as a slight end effect which would be very difficult to describe quantitatively.

It can also be readily shown from Eq. (4) that $G'=-(\omega/a')dX_M/dL$. Thus $G'$ is obtained from the slope of a plot of $X_M$ against $L$ for constant (but not necessarily known) mass. Points corresponding to constant mass are joined by dotted lines in Fig. 4. In this case the intercept may be substantial because of the contributions to $X_M$ from inertia and the elastance of the support, which are independent of immersion (provided,

as must be assumed, the inertia of liquid clinging to the rod is negligible). In Model 1, $r_i = 0.159$ cm, $r_a = 0.285$, and $a' = 110$; in Model 2, $r_i = 0.159$, $r_a = 0.279$, and $a' = 121$. In both models, approximately, $r_0 = 0.18$ cm and $r_1 = 0.23$ cm.

## ILLUSTRATIVE RESULTS

Only a few illustrative data will be given here, reserving systematic studies, correlations with other methods, and interpretations of the data for subsequent papers.

### Behavior of Liquids of Low Molecular Weight

For a liquid which does not contain polymer molecules, there should be no distinction between $\eta'$ at audio frequencies and $\eta$, the viscosity in steady flow. Accordingly, measurements were made on a sample of honey and two viscous oils furnished by the Bureau of Standards. Representative plots of $R_M$ against $L$ are shown in Fig. 5. Values of $\eta'$ calculated from the slopes of such lines are compared in Table I with the steady flow viscosities as provided by the Bureau of Standards or measured by the falling ball method. The agreement is satisfactory except for oil OB-1, which may not be a simple low molecular weight liquid inasmuch as its viscosity slowly increases with time. The indicated accuracy is adequate for preliminary study of the phenomena observed in polymer solutions as illustrated below.

### Behavior of Concentrated Polymer Solutions

The concentrated solutions of linear polymers thus far investigated fall roughly into two classes: those in which $\eta'$ is of the same order of magnitude as $\eta$, but falls off with increasing frequency; and those in which $\eta'$ is very much smaller than $\eta$.

The first type of behavior is illustrated in Fig. 7 for a 20.2 percent solution of polyvinyl acetate (unfractionated, number-average molecular weight[13] 128,000) in cyclohexanone. The steady flow viscosity (falling ball) is 21.8 poises; the dynamic viscosity is nearly equal to this at 32 cycles/sec., but falls rapidly with increasing frequency. At the same time, the dynamic rigidity,

[13] We are indebted to Mr. G. V. Browning for this value. The polymer (AYAX, lot 1232) was kindly furnished by Mr. A. K. Doolittle, Carbide and Carbon Chemicals Corporation.

TABLE I. Comparisons of dynamic and steady flow viscosities of ordinary liquids at 25°C. Dynamic viscosities calculated from measurements in closed tube (Eq. 6) except where otherwise indicated.

| Liquid | Transducer model | Frequency cycles/sec. | $\eta'$ poises | $\eta$ poises |
|---|---|---|---|---|
| Oil N-13 | 1 | 40 | 10.4 | 8.9[a] |
| | 1 | 160 | 11.4 | |
| | 2 | 200 | 9.2 | |
| | 2 | 250 | 9.4 | |
| | 2 | 320 | 9.5 | |
| | 2 | 400 | 9.2 | |
| Honey | 2 | 320 | 49.3 | 51.1[b] |
| Oil OB-1 | 2 | 320 | 195 | 254[b] |
| | | | 199[c] | |

[a] Value given by Bureau of Standards.
[b] Determined by falling ball method.
[c] Calculated from measurements in U-tube at deep immersions (Eq. 5).

which is by definition zero in steady flow, is extremely small at 32 cycles/sec., but rises rapidly with increasing frequency and attains a value of $1.15 \times 10^4$ dyn/cm$^2$ at 400 cycles/sec. These changes represent a striking example of the transition from liquid-like to solid-like behavior.

The second type of behavior is illustrated in Fig. 8 for solutions of polyisobutylene (unfractionated, viscosity-average molecular weight[14] 1,200,000) in xylene. The dynamic viscosity is 30 to 500 times smaller than the steady flow viscosity; their ratio is smallest for the most concentrated solution, and it falls off with increasing frequency as before. The dynamic rigidi-

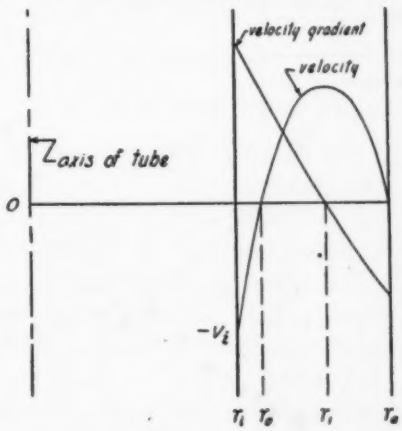



FIG. 6. Schematic plot of velocity and velocity gradient between rod and tube.

[14] J. N. Ashworth and J. D. Ferry, J. Am. Chem. Soc. (in press). The polymer (303-92-1) was kindly furnished by Dr. John Rehner, Jr., Esso Laboratories.

TABLE II. Results of transducer measurements on solutions of polyisobutylene in xylene at 25°C; comparison with data from transverse wave propagation (see reference 14).

| Conc. wt. % | $\eta$ poises | $G'$ Transducer dyn/cm² ×10⁻⁴ | $G'$ Waves[a] dyn/cm² ×10⁻⁴ | $\eta'$ Transducer[b] poises | $\eta'$ Waves[a,b] poises |
|---|---|---|---|---|---|
| 10.5 | 145 | 0.65 | 0.63 | 2.6 | >1.0 |
| 12.5 | 429 | 1.02 | 1.03 | 3.3 | >1.7 |
| 16.7 | 1425 | 2.74 | 2.48 | 5.5 | 4.1 |
| 18.6 | 2780 | — | 3.26 | 6.1 | 5.3 |

[a] Calculated assuming a plane wave in an infinite medium,[1] with $\lambda/x_0$, the damping index, estimated as 1.0.
[b] At a frequency of 320 cycles/sec.

ties of these solutions are sensibly independent of frequency; they are summarized in Table II and shown to be in satisfactory agreement with values derived from wave propagation measurements.[14] The values for steady flow viscosity are also given in Table II, together with rough estimates of the dynamic viscosity obtained from the damping of transverse waves, which, considering the uncertainty inherent in damping measurements, again confirm the transducer results quite satisfactorily. The concurrence of the two entirely different methods is gratifying.

### Significance of the Behavior of Polymer Solutions

The only data in the literature with which Figs. 7 and 8 can be compared are those of Philippoff[4] for solutions of cellulose acetate in dioxane; in his more dilute solutions the behavior of $\eta'$ at audio frequencies resembled Fig. 7, and in his more concentrated solutions, Fig. 8.

Qualitatively, the Maxwell element with spring ($G$) and dashpot ($\eta$) in series[1,3] explains all these results. For this mechanical model, in the frequency range where $\omega \sim G/\eta$, with increasing frequency $\eta'$ falls from $\eta$ toward 0 and $G'$ rises from 0 toward $G$, as in Fig. 7. In the frequency range where $\omega \gg G/\eta$, $G'$ is almost equal to $G$, and $\eta' \ll \eta$; $\eta'$ still decreases with increasing frequency, as in Fig. 8. Probably the only distinction between the two types of behavior cited is that in the first there is a principal zone of dispersion at audiofrequencies, while in the second the dispersion zone lies well below the frequency range covered experimentally. With increasing concentration, the dispersion is shifted to lower frequencies. For a quantitative fitting of these data, however, the Maxwell element is inadequate, and more complicated models must be employed.[1,3]

From a molecular standpoint, the rigidity may be attributed to a transient network structure formed by temporary entanglements of the linear molecules, in some cases perhaps enhanced by associative forces.[15] The frequency range in which the principal dispersion zone falls, corresponding to the transition between elastic deformation of the network structure ($G'$ substantial, $\eta'$ small) and steady flow ($G'$ zero, $\eta'$ large), will be determined by a number of factors, including the extent of overlapping of the individual polymer coils and their rates of diffusion relative to each other.

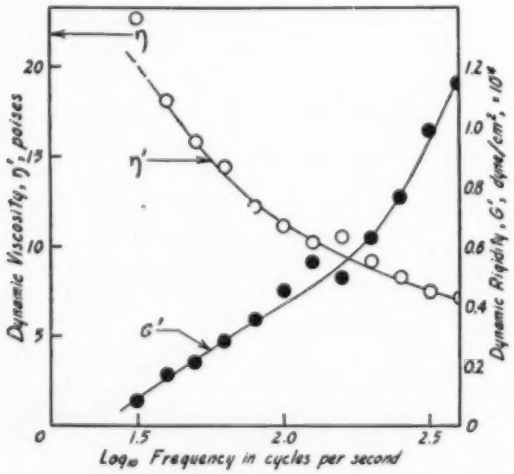



FIG. 7. Frequency dependence of $\eta'$ and $G'$ for 20.2 percent polyvinyl acetate in cyclohexanone at 35°C.

## CRITIQUE OF METHOD

### Range of Frequencies and Consistencies Covered

The transducers can be used over frequency ranges not too far removed from their respective mechanical resonance frequencies. The resonance frequency is $\nu_0 = (S_M/M)^{\frac{1}{2}}/2\pi$, where $S_M = S_M^0 + a'LG'$; the values in air ($G'=0$) are 10 cycles/sec. for Model 1 and 400 cycles/sec. for Model 2. The presence of a sample with substantial rigidity $G'$ does not change the resonance frequency of Model 2 very much, but may raise that of Model 1 to as high as 150 cycles/sec.

Measurements at the exact resonance frequency are to be avoided because the amplitude

[15] J. D. Ferry, Bur. Stand. J. Research 41, 53 (1948).

of vibration may be excessive and the bridge balance somewhat unstable. On the other hand, at frequencies too far from the resonance frequency the difference $R-R_0$ appearing in Eqs. (1) and (2) becomes too small to measure accurately. The breadth of the band of usable frequencies is approximately proportional to $R_M$ and inversely proportional to $(MS_M)^{\frac{1}{2}}$. Thus it is always advantageous to keep the mass small. It is better to have $S_M$ also small, but in an apparatus designed for high frequencies this is not possible; thus, for Model 2 the usable frequency range is considerably smaller (on a logarithmic scale) than for Model 1. From the standpoint of the breadth of frequency band it is better to have $R_M$ large, but from the standpoint of the magnitude of the difference $R-R_0$ near the center of the band it is better to have $R_M$ small (Eq. (1)). It follows that liquids of low viscosity can be measured more accurately but over a narrower frequency range than liquids of high viscosity. For practical purposes, the frequency limits have been 30–200 cycles/sec. for Model 1 and 200–500 cycles/sec. for Model 2; the range of values of $\eta'$ which have been measured is roughly 1 to 200 poises; the range of values of $G'$ is roughly $10^3$ to $10^5$ dyn/cm$^2$ for Model 1, and $10^4$ to $10^5$ dyn/cm$^2$ for Model 2.

### Magnitudes of Displacement, Shear, Velocity, and Rate of Shear

The maximum vertical displacement of the rod and coil is $|x|=\sqrt{2}BlI\times10^{-1}/\omega|\mathbf{Z_M}|$, where $I$ is the r.m.s. current through the coil; the maximum velocity is $|\dot{x}|=\sqrt{2}BlI\times10^{-1}/|\mathbf{Z_M}|$. In typical experiments with $I=0.01$ amp., $|x|$ varies from $4\times10^{-4}$ cm to $10^{-2}$ cm; $|\dot{x}|$ varies from 0.3 to 6 cm/sec. By decreasing $I$ when $|\mathbf{Z_M}|$ is decreased (at low viscosities or shallow immersion) the velocity can be kept below 1 cm/sec. if desired.

The shear, $\partial x/\partial r$, and rate of shear, $\partial\dot{x}/\partial r=\partial v/\partial r$, are greatest at the surface of the rod ($r=r_i$). From Eq. (7) it follows that

$$(\partial v/\partial r)_{r=r_i}=v_i(r_a^2-r_i^2)/r_i[(r_a^2+r_i^2)\times\ln(r_a/r_i)-(r_a^2-r_i^2)]=Kv_i.$$

The constant $K$ is about 60 cm$^{-1}$ for both transducers. Noting that $\partial x/\partial r=(\partial\dot{x}/\partial r)/\omega$, it is found by setting $v_i=|\dot{x}|$ that the maximum shear in typical experiments varies from 0.02 to 0.1 and the maximum rate of shear, or velocity gradient, from 20 to 300 sec.$^{-1}$. They can be kept near the lowest of these respective values if desired.

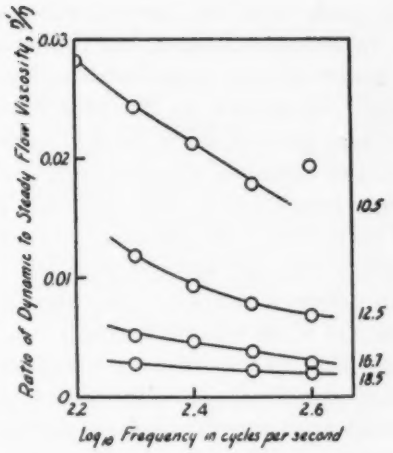



FIG. 8. Frequency dependence of $\eta'/\eta$ for solutions of polyisobutylene in xylene at 25°C. The figures opposite the curves denote concentrations in weight percent.

The shear values are quite small, and there is no reason to expect non-Hookean elasticity due to excessive distortion of the molecular entanglement. The rates of shear, although not extremely high, are in the range where non-Newtonian effects appear in concentrated polymer solutions in steady flow. However, it is likely that in small oscillating deformations higher rates of shear can be tolerated than in steady flow without encountering a non-linear relation between stress and rate of shear; the stresses involved are smaller because $\eta'<\eta$. Evidence for this viewpoint follows.

### Absence of Non-Linear Effects

If enormous complications in the interpretation of data are to be avoided, experimental conditions must fulfil as nearly as possible the requirement that the real and imaginary parts of the driving force be linear in the strain and rate of strain respectively, *i.e.*, that $G'$ and $\eta'$ be independent of the amplitude of vibration at constant frequency.

In several cases this requirement has been tested by varying the driving current, and hence the amplitude, at constant frequency. In one series with oil N-13, measurements were made at 320 cycles/sec. with transducer Model 2 at driving currents of 0.0025, 0.005, 0.01, and 0.015

amp. For each value of current, the amplitude actually varied depending on the immersion and on the amounts of added mass used in constructing circle plots. Nevertheless, the four circle plots and the linear plots of $R_M$ against $L$ were almost identical; the values of $\eta'$ calculated from these data were, respectively, 9.7, 9.7, 9.5, and 9.5 poises. This indicates that there are no non-linear effects inherent in the apparatus. A few isolated tests with polymer solutions have also shown no dependence of $\eta'$ on driving current.

The fact that in measurements with constant current the amplitude actually varies with immersion and mass permits additional evidence to be adduced for the constancy of $G'$ and $\eta'$. If these quantities depended on amplitude, the plots of $X$ against $R$ in Method 2 would not be circular, and, moreover, the plots of $R_M$ and $X_M$ against $L$ would not be linear. It is concluded that within the accuracy of the method we are dealing with linear viscoelastic behavior[1,3] in the systems studied.

### Neglect of Inertia of the Sample

In the mechanical and hydrodynamic calculations no account has been taken of the inertia of the sample. There are two possible sources of error in this neglect: (1) some of the solution, clinging to the rod, may add to its mass so that the inertia contribution to $X_M$ will not be independent of immersion, thus introducing an error in the calculation of $G'$ from a plot of $X_M$ against $L$; (2) if the inertia of the solution is sufficient to set up an incipient transverse wave in the gap between rod and tube, the hydrodynamic calculations given will not be adequate.

The relative magnitude of the first error is gauged by comparing $\omega dM/dL$ with $a'G'/\omega$, where $dM/dL$ is the increase in effective mass with immersion due to clinging of liquid. If the entire cylindrical shell which moves in the same direction as the rod, between $r_i$ and $r_0$ (Fig. 6), is assumed to add to the inertia, $dM/dL \cong 0.02$ g/cm; this is probably an overestimate. Such a value for $dM/dL$ could introduce a negative error of 10 percent in $G'$ at 400 cycles when $G'$ is about $10^4$ dyn/cm$^2$. The agreement of $G'$ values calculated from transducer and wave measurements in a number of cases (e.g., Table II) indicates that the error is no greater than this. Since the relative error is proportional to the square of the frequency, it may be ignored at low frequencies. However, it may set an upper limit to the usable frequency range for calculations of $G'$. Fortunately, it does not affect the calculation of $\eta'$.

The second source of error can be disregarded if the thickness of the sample subjected to shear is less than $(\eta'/\rho\omega)^{\frac{1}{2}}$ or less than $(G'/\rho)^{\frac{1}{2}}/2\omega$, as pointed out elsewhere.[1] For this calculation, the sample may be regarded as two shells of thickness $r_a - r_1 = 0.05$ cm and $r_1 - r_i = 0.07$ cm which are being sheared in opposite directions (Fig. 6). In the polyvinyl acetate solutions (Fig. 7) at low frequencies $(\eta'/\rho\omega)^{\frac{1}{2}}$ is about 0.3, so the first inequality is satisfied. At high frequencies both $(\eta'/\rho\omega)^{\frac{1}{2}}$ and $(G'/\rho)^{\frac{1}{2}}/2\omega$ are about 0.03, so neither equality is quite satisfied; a rough calculation indicates that in this case both $\eta'$ and $G'$ will appear to be somewhat too large.

### Comparisons with Other Methods

In certain cases different methods for measuring $G'$ and $\eta'$ of concentrated polymer solutions may be preferable to our transducers; while in other cases the transducers have advantages not possessed by any other method. In contrast to the Philippoff apparatus,[4] the transducers provide measurements of $G'$ as well as of $\eta'$; measurements can be made over a range of frequencies, spaced as closely as desired, without mechanical adjustments to change the resonance frequency; the amplitude of oscillation is smaller; and the electrical circuit is simpler. However, the values of $\eta'$ obtained appear to be somewhat less precise than those of Philippoff. In comparison with the methods of Goldberg and Sandvik[16] and Van Wazer and Goldberg,[17] which, like Philippoff's, depend on resonance effects, the transducers cover a different frequency range; the Goldberg devices are adapted to the range from 5 to 50 cycles/sec. They are inoperative if the mechanical loss factor $(\omega\eta'/G')$ is too high. In comparison with the wave propagation method,[2] the transducers give much more reliable measurements of $\eta'$, and they provide measurements of both $\eta'$ and $G'$ below the principal zone of dispersion where

[16] H. Goldberg and O. Sandvik, Anal. Chem. **19**, 123 (1947).
[17] J. R. Van Wazer and H. Goldberg, J. App. Phys. **18**, 207 (1947).

wave propagation fails because of extreme damping. However, transducer measurements of $G'$ above the principal zone of dispersion are less reliable than wave measurements. The transducer and wave methods complement each other quite satisfactorily, and by combining them it should be possible to determine both $G'$ and $\eta'$ for polymer solutions over a considerable frequency range.


### ACKNOWLEDGMENT

This work was supported in part by the Research Committee of the Graduate School of the University of Wisconsin from funds supplied by the Wisconsin Alumni Research Foundation, and in part by a grant from Research Corporation. We are much indebted to the mechanicians in the Department of Chemistry, especially to Mr. Lloyd Lincoln, for machine work in construction of the apparatus, and to Mr. Myrle F. Johnson and Mr. George V. Browning for help in making some of the measurements.


### APPENDIX

## Mechanical Resistance of a Rod Moving Axially in a Closed Tube of Viscous Liquid[18]

A rod of radius $r_i$ is immersed to a depth $L$ in a liquid of viscosity $\eta$ contained in a tube of radius $r_a$, closed at the bottom. When a force $f$ is applied to the rod, it will move down with velocity $-v_i$ (negative because it is opposite to the net movement of liquid (Fig. 6)). The velocity of the liquid, $v$, is described by the general integral for Poiseuille flow:[19]

$$v=-pr^2/4\eta L+(A/2\pi\eta L)\ln r+B,$$

where $p$ is the increase in pressure at the bottom of the tube which is responsible for the upward displacement. The constants $A$ and $B$ are determined by our boundary conditions: $v=0$ at $r=r_a$ and $v=-v_i$ at $r=r_i$. Thus

$$v=(p/4\eta L)(r_a^2-r^2)+[\ln(r/r_a)]\times[v_i+(p/4\eta L)(r_a^2-r_i^2)]/\ln(r_a/r_i). \quad (9)$$

Since the volume of liquid displaced by the end of the rod must be equal to the volume forced up around the rod, $v_i\pi r_i^2=\int_{r_i}^{r_a} v\cdot 2\pi r dr$. Evaluating and solving for $p$,

$$p=4\eta L v_i/[(r_a^2+r_i^2)\ln(r_a/r_i)-(r_a^2-r_i^2)]. \quad (10)$$

By substituting (10) in (9) and differentiating with respect to $r$ we obtain the equations previously cited, (6) and (7). The force $f$ is equal to the viscous drag of the adjacent liquid plus the hydrostatic force on the bottom of the rod. The former is the product of the velocity gradient, the area immersed, and the viscosity; the latter is the product of the pressure and the cross section area of the rod:

$$f=2\pi r_i L\eta(\partial v/\partial r)_{r=r_i}+\pi r_i^2 p. \quad (11)$$

From (7) with $r=r_i$ and (10) we find

$$f=\frac{2\pi\eta L v_i(r_a^2-r_i^2)}{[(r_a^2+r_i^2)\ln(r_a/r_i)-(r_a^2-r_i^2)]\ln(r_a/r_i)}+\frac{2\pi\eta L v_i}{\ln(r_a/r_i)}.$$

By definition, $R_M=f/v_i$; hence, with rearrangement,

$$R_M=\frac{2\pi\eta L(r_a^2+r_i^2)/(r_a^2-r_i^2)}{(r_a^2+r_i^2)\ln(r_a/r_i)/(r_a^2-r_i^2)-1}=a'\eta L.$$

For cyclic deformations, since $R_M$ is defined as the force in phase with velocity divided by velocity, the above holds with substitution of $\eta'$ for $\eta$ (Eq. (8)).

If the force and viscosity are considered as complex ($\mathbf{n}=\eta'-i\eta''$), the above equation holds with $\mathbf{Z_M}$ and $\mathbf{n}$ substituted for $R_M$ and $\eta$, and it follows from the relation[1] $\eta''=G'/\omega$ that $G'=-(\omega/a')dX_M/dL$.

A somewhat similar problem has very recently been treated by Bikerman.[20]

[18] This problem was treated by Lawaczeck (Z. Ver. deutscher Ing. **43**, 677 (1919)) for the approximation of a very narrow gap between rod and tube; his approximation is inadequate for our purposes, however.

[19] Page, *Theoretical Physics* (D. Van Nostrand and Company, Inc., New York, 1935), p. 259.

[20] J. J. Bikerman, J. Colloid Sci. **3**, 75 (1948).

# Density and Packing in an Aggregate of Mixed Spheres

DOUGLAS RENNIE HUDSON
*Higher Teachers' College, Baghdad, Iraq*
(Received July 3, 1947)

It is shown that in close packing of spheres, two types of interstice exist, bounded by six and by four convex spherical surfaces. These are termed "square" and "triangular." They are connected by a continuous labyrinth through which a ball not exceeding $(2/\sqrt{3}-1)r$ in radius can be threaded. In both cubic and hexagonal arrangements, their shape and size are identical, but their distribution differs. Radii $R$ for $n$ smaller balls, which can take up patterns with cubic symmetry within each square interstice, are calculated for values of $n$ up to 27 and plotted. The expression

$$(R/r)=(\sqrt{2}-1)/n^{\frac{1}{3}}$$

is used as a test for efficiency of packing.

When $R/r$ is plotted against the density increment attributable to this interstitial packing, a set of spires is obtained. At $n=8$ and $n=9$ ($R/r$ being, respectively, 0.2289 and 0.2166) twin peaks of 13 percent and 12 percent appear, which are accentuated by first entry of a sphere into the triangular interstice at $R/r=0.22475$; this contributes another 3 percent. The peak of 16 percent at $n=21$, $R/r=0.1782$, is also reinforced by first entry of 4 spheres in a body-centered tetrahedron at 0.1716, which gives an additional 6 percent density increment. Applications to bulk storage, ceramics, and interstitial compounds and solid solutions are considered. None of the special packs gives a density increment $(\partial\Delta_{\text{square}}+\partial\Delta_{\text{triangular}})/\Delta$ even approaching the 26 percent for fine spherical filler close packed in the interstices.

APART altogether from theoretical interest in the study of metallic solid solutions,[1] packing of unequal spheres is important in handling industrial substances, e.g., a mixture of two sizes of oranges, or sago, or nickel shot. Recently, considerable interest has been taken in the density of aggregates of less regular substances, for instance, coal and iron ore.[2–4] Less directly, the question is of importance in industrial sieving and grading, ore dressing, concrete technology and soil physics,[5, 6, 7] and in all processes of grinding by attrition, whether they concern cement in a tube mill, auriferous quartzite on the Rand, or silver ore in a primitive Mexican *arrastra*.

The present communication summarizes a theoretical investigation of smaller spheres introduced into interstices of an array of spheres in close packing. Very interesting results have been obtained for the radius ratio, the number that can be introduced, and the consequent rise in the density of the aggregate $\partial\Delta/\Delta$ equal to $n(R/r)^3$.

## 1. COORDINATION NUMBER

Regular close packing of spheres is possible in two ways, cubic and hexagonal, and four mixtures of these, but the number of types of true close packings is unlimited. Although the unit cells (the bricks of which they are composed) are not identical, they are very similar. In each the coordination number, or number of spheres in contact with any given sphere, is *twelve*. This is familiar to everybody, but it is not so generally appreciated that *these twelve are not, in fact, distributed regularly around the central sphere.*[1, 3, 4, 8, 9] Figures 1A and 1B are photographs of the same close packing from two viewpoints. *Interstices between the spheres are bounded alternately by six and by four spherical surfaces. By analogy from spherical trigonometry these might be termed reversed spherical cubes and reversed spherical tetrahedra.* We may conveniently, if incorrectly, call them "square" and "triangular" holes. With every sphere, one square and two triangular holes are associated. Whether the original close packing is cubic or hexagonal or a mixed packing

[1] D. R. Hudson, J. App. Phys. **17**, 1126 (1946).
[2] H. T. Horsfield, J. Soc. Chem. Ind. **53**, 107T (1934).
[3] R. L. Brown and P. G. W. Hawksley, Nature **156**, 421 (1945).
[4] D. R. Hudson, Machinery **70**, 617, (1947).
[5] R. R. Litehiser, "Effect of moisture in sands in proportioning concrete mixtures," Concrete (1925) (not seen).
[6] A. E. R. Westman and H. R. Hugill, J. Am. Cer. Soc. **13**, 767 (1930).
[7] H. E. White and S. F. Walton, J. Am. Cer. Soc. **20**, 155 (1937).
[8] D. R. Hudson, Proc. Leeds Phil. Soc. **5**, 73 (1947–48).
[9] L. C. Graton and H. J. Fraser, J. Geology **43**, 785 (1935).

merely alters the distribution of the two types, not their nature or size.

"We proceed to insert, in turn, 1, 2, 3, 4, $5\cdots n$ equal smaller balls into each *square* hole, and to calculate the radius $R$ in terms of the radius $r$ of the spheres in the original array ($R<r$), by straightforward, if tedious, solid geometry (Fig. 5). From Figs. 1A and 1B the triangular interstice is much the smaller, so that in the first place we need not consider it. But as the number of balls in each square hole increases, and their diameter decreases, we eventually reach at $0.22475r$ a size small enough for a single ball to be inserted in each triangular hole. So long as $R$ exceeds this, the triangular holes can contribute no increment to the density.

### 2. EFFECTIVE DENSITY OF THE PACK

The effective density of the original close packing is the ratio:

$$\frac{\text{volume of spheres}}{\text{total volume}} = \frac{\text{volume of spheres}}{(\text{volume of spheres}+\text{volume of interstices})}.$$

The volume of a sphere is $4\pi r^3/3$, and that of the space associated with each $8r^3/3$, giving an effective density of $\pi/3\sqrt{2}$ or 0.7405.

### 3. SYMMETRICAL PACKING WITHIN A SQUARE INTERSTICE

#### Density Increments for Values of $n$ up to 9

$n=1$: For a single ball in each square hole $R=(\sqrt{2}-1)r=0.4142r$ almost by inspection, and the increment in density $\partial\Delta/\Delta$ is obviously $nR^3/r^3$ or $(0.4142)^3$, i.e., 0.07106. To the eye this packing appears good.

$n=2$: For two balls in each square hole—lying along the cube diagonal—$R$, from the equation $2R(\sqrt{2/3}+1)=r$, is found to be $0.2753r$. The poverty of this pack, apparent on visual examination, is confirmed on calculation, the increment in density being only 0.04170.

$n=3$: No symmetrical arrangement is possible. Though a formal proof is elusive, it is believed that three balls cannot be inserted without reducing the radius to such a value that four can enter.

$n=4$: This pack also appears poor. To give maximum symmetry the balls must lie along diagonals in parallel faces, crossed at right angles. From the equation $r^2-4rR+R^2/2=0$, $R=0.2583r$, and the increment in density is 0.06896. Despite apparently poor packing, it is very nearly the same as that given by a single ball in each square interstice.

$n=5$: No symmetrical packing is possible that will not admit a larger number of balls.

$n=6$: This case is very interesting. By inspection $R=(\sqrt{2}-1)r/(\sqrt{2}+1)$ or $0.17157r$. This surprisingly low value has no meaning, since with this radius a larger number of balls can be packed into the square hole. A later calculation will show us that this is as high as 21 at least.

$n=7$: For any symmetrical packing the remarks under "$n=6$" apply with even greater emphasis.

$n=8$: This is obviously the best pack yet made. The balls tuck into the open corners while the void at the center remains small. From the equation $r^2-2rR(\sqrt{2}+1)+2R^2=0$, $R=0.22885r$, and the density increment of 0.09590, the highest yet obtained, confirms our expectations.

$n=9$: This pack is a modification of the former, with an additional ball at the center, each corner ball being displaced outwards diagonally by the small amount necessary to accommodate it. To the eye the pack appears fairly close and the central void is filled by the new introduction. Thus, *a priori*, one might expect an increment slightly larger than that for $n=8$. From $r^2-2rR(2\sqrt{2/3}+1)+3R^2=0$, $R$ is found to be $0.21663r$. The corresponding increment in density is $9\times(0.21663)^3$ or 0.09150. This pack corre-

(A) "Square" interstice. (B) "Triangular" interstice.

FIG. 1. Showing two layers of a Barlow close packing. Until the third layer is added it is impossible to say whether cubic or hexagonal close packing will develop.

TABLE I. Showing the limiting (smallest) radius for $3^3$ and derived packs within a square interstice.

| Location | $R/r$ for $3^3$ | $R/r$ for $(3^3-1)$ | $R/r$ for $(3^3-6)$ |
|---|---|---|---|
| Corner | 0.1742 | Slightly larger than in $3^3$ | 0.1782 |
| Center of face | 0.1516 | 0.1547 | 0.1833 |
| Center of edge | 0.13807 | 0.1716 | Removed |

sponds to the unit cell of the body-centered cube as usually drawn.

### Density Increments for $n=14$, 16, and 17

For $n=10$, 11, 12, and 13, we have been unable to devise packs with cubic symmetry which will not accommodate a greater number.

$n=14$: This corresponds exactly with the unit cell for cubic close packing as usually drawn; one would expect it to give a good density increment. Four corner balls and the one lying midway between them have centers in one plane. The bounding sphere will obviously touch the ball in the center of the face just where its own bulge is greatest (the corresponding value of $R$ is $0.1716r$, as with six balls), but will not touch corner balls. *We shall define such a pack, in which the bounding spheres are not in contact with all interstitial balls, as slack, and the case where they are in contact with sufficient balls to present movement, as tight.* A *slack* aggregate would rattle on shaking (like a Cox's Orange Pippin) and the term has *no connection with "close packing."* In this particular pack the point is very easy to grasp, and the density increment is $(0.1716)^3\times 14$ or 0.07074. In more complicated arrangements it may not be so easy to eliminate contacts by inspection; one must then calculate values of $R$ for various positions. Where these are not equal (as will generally be the case) the lowest must be taken since the pack is *slack*.

$n=15$: In this pack, corresponding to the former with an additional ball at the center, $R$ is reduced to $(\sqrt{2}-1)r/3$, or $0.13807r$. Even with simple cubic packing, 27 balls of this size can be introduced, and with re-arrangement many more. The case is analogous with that of $n=6$.

$n=16$: Two concentric sets of eight balls. Although the original eight gave such a good density increament (0.09590) this double-eight is not nearly so promising to the eye in view of the large vacant space. From the equation $r^2-2rR(1+2+2\sqrt{2/3})+R^2(6+4\sqrt{3})=0$, $R=0.16933r$, giving a density increment of 0.07768.

$n=17$: Here the pack is similar, with an extra ball introduced into the central void, and corresponding displacement of other balls diagonally outward. From the equation $r^2-2rR(1+4\sqrt{2/3})+15R^2=0$, $R=0.16517r$, and the corresponding increment is 0.07660.

### Density Increments for $n=27$, 26, and 21

$n=27$: The next primitive cube has $3^3$ balls, $(R/r)$ is $(\sqrt{2}-1)/3$ or 0.13807; packing is obviously very loose as is confirmed in the low value of 0.07108 for $\partial\Delta/\Delta$. In this arrangement, as with $n=14$, contact is obviously between the ball at the center of each face and the bulge of the spherical bounding surfaces. In Table I $(R/r)$ is calculated for contact of the bounding spheres with balls in the following positions in the $3^3$ arrangement:

(*a*) corner,
(*b*) center of edge,
(*c*) center of face.

Values of $(R/r)$ decrease in this order; if the balls be considered fixed, a bounding sphere drawn in contact with four corner balls would cut those in the centers of the edges and the centers of the faces.

$n=26$: We should get a much better pack if each face were "dished" by removing a single ball from the center of the cube, and pushing inward that in the middle of each face, to fill the space. The result is a considerable increase in $(R/r)$ for the ball at the center of the face, from 0.1381 to 0.1716; clearly this is no longer the limiting ball. The radius of the corner ball is hardly affected and that at the center of the edge not very much, $(R/r)$ being changed from 0.1516 to 0.1547. This is now the limiting factor, since the packing is *slack*. The corresponding increment in density is $26(0.1547)^3$ or 0.09626.

Consideration shows that a pack with good density increment will have the following characteristics:

(*a*) dished faces,
(*b*) balls tucked into the spacious corners where the convex spherical surfaces sweep back.

Physical intuition might have told us as much, but how deceptive this may be!

These conditions are satisfied by the good packs given by 8, 9, and 26 balls, as shown by the high density increments of 0.09590, 0.09150 and 0.09626, respectively. There is actually a third condition—that the arrangement within the interstice shall approximate to close packing, but in our treatment there are not enough even to approach this.

$n=21$: An alternative variation is to remove one ball from the center of each face and allow the remainder to accommodate themselves in a hopper-faced cube. We are not then concerned with contact between the spherical boundaries and the ball in the center of each face, but must compute $(R/r)$ separately for the ball at the corner, and at the center of each edge. These are found, respectively, to be 0.17824 and 0.1835, so that the density increment is $21(0.17824)^3$, or 0.11892. This is the best increment we have yet attained for interstitial packing in the square holes.

Horsfield[2] has investigated theoretically the effect of inserting into a close-packed array successively smaller spheres just large enough to fill the interstices. Table II is calculated from his results.

Thus, by using spheres of five sizes, an effective density 115 percent of that of the original close-packed aggregate can be attained. This very complicated pack has a voidage of 14.9 percent. If its interstices were filled with material fine enough to be in close packing, a further density increase of (14.9 percent$\times$0.76) or 11.3 percent would be possible—a total increment of 26.2 percent with corresponding voidage of 3.9 percent.

## 4. SYMMETRICAL PACKING WITHIN A TRIANGULAR INTERSTICE

$n=1$: The largest ball able to enter a triangular hole has a radius $0.22475r$. Since there are two such interstices for every sphere in the aggregate, a density increment of $1(0.22475)^3\times2$, or 0.0227, is attributable to them. At this value, eight balls can enter every square interstice in a slack degenerate $2^3$ pack, giving an increment of $8(0.22475)^3$ or 0.09083. Since density increments are additive, it is better to consider separately that due to triangular holes.

$n=4$: As $R/r$ decreases, the increment in density diminishes until $R/r=0.1716$, at which four balls can enter each triangular hole. Consequently, at this point there is a fourfold rise in the increment to 0.040424. The packing in the square hole accommodating twenty-one balls in hopper-faced cube arrangement has $(R/r)=0.1782$. At the point where four balls enter the triangular hole, the density increment contributed by the square hole is therefore $21(0.1716)^3=0.0106113$.

$n=5$: We might term this a body-centered tetrahedral pack. $R/r$ is found to be 0.14208, and the corresponding increment from the triangular interstice is $5(0.14208)^3\times2$ or 0.02868.

For a greater number of balls other patterns could be devised, but it does not seem worth while to go much further, especially as 0.14208 is very near the value 0.13807 for $3^3$ balls in simple cube formation, at which we left off calculation for a square interstice.

TABLE II. Density increments and voidages theoretically attainable in complex packs (after Horsfield's data).

| No. of grades | Type of pack | $R/r$ | $N_R$ | Increment in density | Cumulative increment in density | Voidage in mixed aggregate |
|---|---|---|---|---|---|---|
| 1 | Original array | 1 | 0 | — | — | 26% |
| 2 | 1 sphere in sq. interstice | 0.4142 | 1 | 0.07106 | 0.07106 | 20.7% |
| 3 | 1 sphere in triangular interstice | 0.22475 | 2 | 0.02272 | 0.09378 | 19.0% |
| 4 | 8 spheres in pockets surrounding sphere in square interstice | 0.175 | 8 | 0.04288 | 0.13666 | 15.8% |
| 5 | 4 spheres in pockets surrounding sphere in triangular interstice | 0.117 | 8 | 0.01281 | 0.14947 | 14.9% |

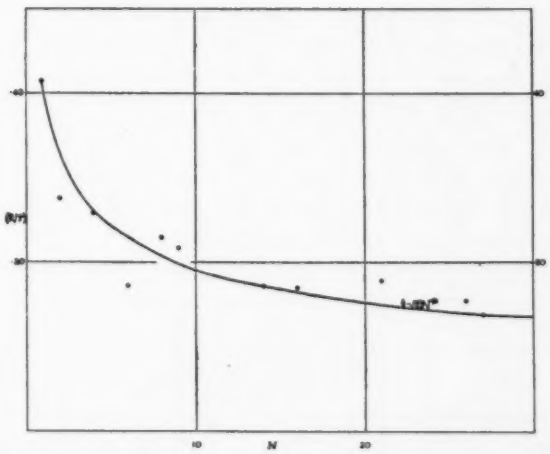


FIG. 2. Relation between the radius ratio $(R/r)$ and the number $N$ of balls of radius $R$ which can be inserted symmetrically into the "square" interstice in an existing close packing of balls of radius $r$. The heavy line $(1-\sqrt{2})/N$ is included for comparison, as a semi-absolute datum.

TABLE III. Density increments for symmetrical packs governed by dimensions of the square interstice.

| Square interstice | | | | | | | | Triangular interstice | | Total density increment |
|---|---|---|---|---|---|---|---|---|---|---|
| Geometric data | | | | | | | | | | |
| $n$ | Stoichiometry | | Type | Equation connecting $R$ and $r$ | $(R/r)$ | Datum for comparison $\sqrt{2}/n^{\frac{1}{3}}$ | Density increment $n(R/r)^3$ | $n$ | Density increment $2n(R/r)^3$ | $\frac{\partial\Delta_{\text{triangular}}}{\Delta}+\frac{\partial\Delta_{\text{square}}}{\Delta}$ |
| 1 | Simple cube | Primitive, $1^3$ | tight | $R=(\sqrt{2}-1)r$ | 0.4142 | 0.4142 | 0.07106 | 0 | — | 0.07106 |
| 2 | Along cube diagonal | | tight | $2R(\sqrt{2/3}+1)=r$ | 0.2753 | 0.3287 | 0.04170 | 0 | — | 0.04170 |
| 4 | Crossed parallel face diagonals | | tight | $r^2-4rR+R^2/2=0$ | 0.2583 | 0.2609 | 0.06896 | 0 | — | 0.06896 |
| 6 | Centers of cube faces | | tight | $R(\sqrt{2}-1)=r(\sqrt{2}+1)$ [See $n=14$, 21] | 0.1716 | 0.2280 | 0.03028 | 4 | 0.04038 | 0.07066 |
| 8 | Simple cube | Primitive, $2^3$ | tight | $r^2-2rR(\sqrt{2}+1)+2R^2=0$ | 0.2288 | 0.2071 | 0.09590 | 0 | — | 0.09590 |
| 9 | Body-centered cube | Primitive, $2^3+1$ | tight | $r^2-2rR(2\sqrt{2/3}+1)+3R^2=0$ | 0.2166 | 0.1992 | 0.09150 | 1 | 0.02034 | 0.11184 |
| 14 | Face-centered cube | Primitive, $2^3+6$ | slack | $R(\sqrt{2}+1)=r(\sqrt{2}-1)$ | 0.1716 | 0.1719 | 0.07074 | 4 | 0.04042 | 0.11116 |
| 16 | Concentric simple cubes | Primitive, $2^3+2^3$ | slack | $r^2-2rR(1+\sqrt{2}+2\sqrt{2/3})+R^2(6+4\sqrt{3})=0$ | 0.1693 | 0.1644 | 0.07768 | 4 | 0.03882 | 0.11647 |
| 17 | Concentric cubes, body-centered | Primitive, $2^3+2^3+1$ | slack | $r^2-2rR(1+4\sqrt{2/3})+15R^2=0$ | 0.1652 | 0.1611 | 0.07660 | 4 | 0.03605 | 0.11265 |
| 21 | Hopper-faced cube, body-centered | $3^3-6$ | slack | $r^2-2rR(11/3)+(29/3)R^2=0$ | 0.1782 | 0.1502 | 0.11892 | 1 | 0.01132 | 0.13025 |
| 26 | Hopper-faced cube, face-centered | $3^3-1$ | slack | $r^2-2rR(2+\sqrt{3})+R^2(3+2\sqrt{3})=0$ | 0.1547 | 0.1398 | 0.09626 | 4 | 0.02962 | 0.12588 |
| 27 | Simple cube | Primitive, $3^3$ | slack | $3R=(\sqrt{2}-1)r$ | 0.13807 | 0.13807 | 0.07108 | 5 | 0.02632 | 0.09740 |

TABLE IV. Density increments for symmetrical packs governed by dimensions of the triangular interstice.

| Square interstice | | | | | Triangular interstice | | | | | | Total density increment |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | Stoichiometry | | Density increment $n(R/r)^3$ | Type | Stoichiometry | Type | Equation connecting $R$ and $r$ | $n$ | $(R/r)$ | Density increment $2n(R/r)^3$ | $\frac{\partial\Delta_{\text{triangular}}}{\Delta}+\frac{\partial\Delta_{\text{square}}}{\Delta}$ |
| 8 | Simple cubic | $2^3$, degenerate | 0.09083 | slack | single | tight | $R=(\sqrt{3/2}-1)r$ | 1 | 0.22475 | 0.02271 | 0.11354 |
| 21 | Hopper-faced cube, body centered | $3^3-6$, degenerate | 0.10611 | slack | tetrahedral | tight | $r^2-6rR+R^2=0$ | 4 | 0.1716 | 0.04042 | 0.14653 |
| 26 | Hopper-faced cube, face centered | $3^3-1$, degenerate | 0.07457 | slack | body-centered tetrahedral | tight | $r^2-2rR(\sqrt{3}+2)+3R^2=0$ | 5 | 0.14208 | 0.02868 | 0.10325 |

## 5. DISCUSSION

### Efficiency of Packing in the Square Interstices

Values of $n$ are plotted against $(R/r)$ in Fig. 2. Since these are necessarily integral, it is not legitimate to draw a curve through the points, or even to join them by straight lines though 5.163 balls per square hole may, of course, have a perfectly clear statistical meaning. For comparison, some standard is necessary. A curve with values of $R/r$ proportional to $n^{-\frac{1}{3}}$ has therefore been drawn through the point for $n=1$. On first sight this should pass through all points for which $n$ is a cube; in fact it will pass only through odd values. For instance, the pack of $2^3$ is much better than would be given by this relation. Odd values of simple cubic packing give only poor density increments.

It would be better if there were some more absolute standard for packing efficiency. In the original close-packed lattice voidage is $(1-\pi/3\sqrt{2})$ or 0.26. If this were wholly filled with spheres fine enough to render boundary effects negligible (the exact converse of what we have so far discussed), the actual increment in density would be $(\pi/3\sqrt{2})(1-\pi/3\sqrt{2})$, or 0.1924, making $\partial\Delta/\Delta$ equal to $(1-\pi 3\sqrt{2})$ or 0.26. Both square and triangular interstices are contributors to this value, and for a more absolute comparison curve in Fig. 2 one needs to know their relative volumes. Since they are interconnected by passages bounded by three convex spherical surfaces in one vast labyrinth, it is quite out of question to delimit a boundary between them. However, one can justify a semi-absolute partition on the basis of relative number of balls for a fairly small value on $n$, a good pack being chosen in each case.

(*a*) When $(R/r)=0.1782$ there are 21 balls in each square hole, when $(R/r)=0.1716$ there are 4 balls in each triangular hole. The ratio of density increments is then

$$\frac{21(.1782)^3}{4(.1716)^3\times 2} \quad \text{or} \quad \frac{0.11886}{0.040424}=2.94.$$

(*b*) When $(R/r)=0.2288$ there are 8 balls in each square hole, when $(R/r)=0.22475$ there is 1 ball in each triangular hole. The ratio of density increments is then

$$\frac{8(.2288)^3}{1(.22475)^3\times 2}=4.22.$$

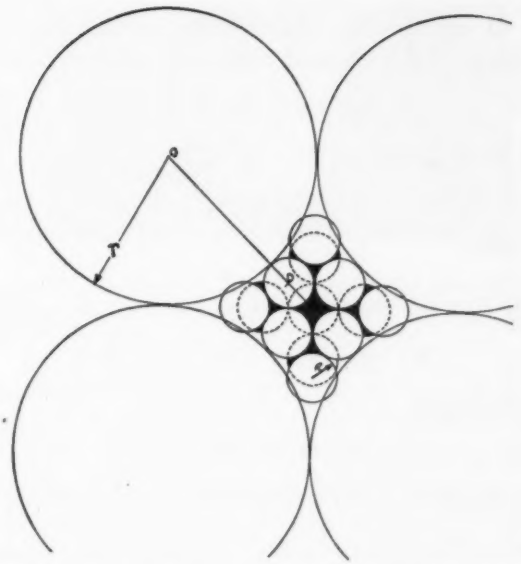


FIG. 3. Pack of 21 balls in a *square* interstice illustrating the calculation of $R/r$ for ball $D$ at the center of the original cube edge. The diagram is for illustration only and does not represent a section through the centers of the larger spheres. The point $B$ is the center of the interstice.

If we take 3 to be a reasonable partition, the increment in density attributable to the square holes is

$$\tfrac{3}{4}\times(1-\pi/3\sqrt{2}), \quad \text{or} \quad 0.195,$$

which gives

$$n(R/r)^3=0.195,$$

i.e., the comparison curve becomes

$$(R/r)=0.580n.$$

Using unit cell planes for delimitation, Graton and Fraser in 1935 calculated a volume ratio of 0.716/0.284, or almost exactly 2.5; if we accept this convention, the factor becomes 0.571. On either basis it is clear that the curve drawn in Fig. 1, through $n=1$, is too low. So far as an absolute comparison is possible, therefore, none of our special packs gives a density increment even approaching that to be expected from close packing of the interstices with very fine spherical filler.

### Computation of $R/r$

Calculation of $R/r$ depends only on intelligent application of Pythagoras' theorem. For such simple geometry one illustration will suffice. Figure 3 shows the pack of $3^3-6$ balls in a *square* interstice, $D$ being the center of the ball at the center of an edge, and $D'$ the projection

of $D$ on the line joining the centers of two spheres of the original pack, of radius $r$.

$$DO^2=D'O^2+D'D^2$$

by Pythagoras. But

$$BD=2R,$$

so that

$$D'B=\sqrt{2}R$$

and

$$(R+r)^2=(\sqrt{2}r-\sqrt{2}R)^2+(\sqrt{2}R)^2,$$

from which

$$r^2-6rR+3R^2=0,$$

$R/r=1\pm\sqrt{2/3}$, the relevant root being 0.1833.

### Peak Densities in the Mixed Aggregate

Data for both square and triangular holes, and the equations from which $(R/r)$ is derived, are summarized in Tables III and IV. There are two well defined peaks in density.

(*a*) The first is due to proximity of good packs of $2^3$ and $(2^3+1)$ at 0.2288 and 0.2166, respectively, giving density increments of 0.0959 and 0.0915. These twin spires are strongly enhanced at 0.2245 by first entry of a single ball into the triangular interstice—the result on the graph in Fig. 4 is very marked.

(*b*) The second peak occurs at $R/r=0.1782$ for twenty-one balls arranged as a hopper-faced cube. Its concave faces accommodate the out-

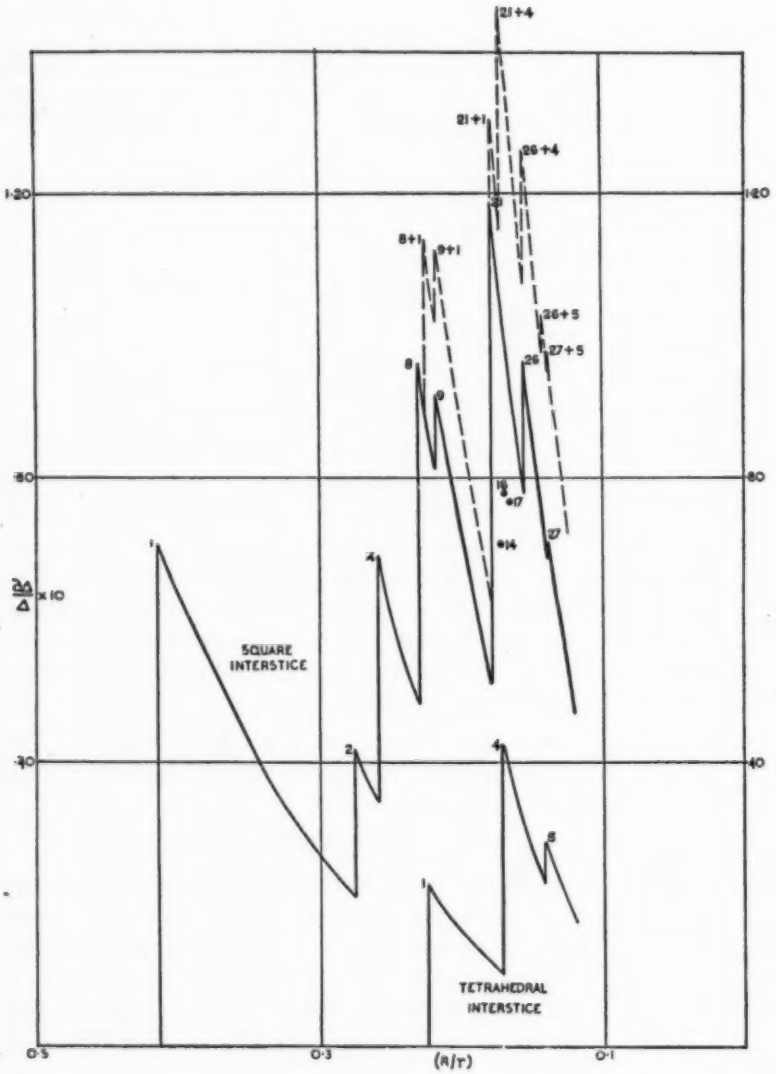



FIG. 4. Relation between the radius ratio $(R/r)$ and the increment in density $(\delta\Delta)/\Delta$ obtained by inserting balls of radius $R$ symmetrically into the "square" insterstices of a close-packed array of balls of radius $r$. Broken line—sum of increments from both types of interstice.

ward bulging spheres; moreover, there is some approach to close packing within the pattern of twenty-one balls. This peak, too, is reinforced, since at $R/r=0.1716$ the number of balls able to enter a triangular hole jumps from *one* to *four*, arranged as a body-centered tetrahedron.

### Application to Commercial Materials

One cannot emphasize too strongly that the computation and graphs apply only to symmetrical packing within the interstices of a close-packed aggregate already existing. It is Mother Nature alone, acting through her interatomic forces, who can pack so regularly, in the unit cell of kaolinite, $\gamma$-brass, feldspar, or complex molybdate. With actual spheres even the original close packing is impossible to attain, despite long-continued jarring.

For lead shot in a cylindrical tube Westman and Hugill in 1930 reported[6] a quite reproducible voidage of 37 percent; this was confirmed[7] by White and Walton in 1937. For dry sand Litehiser in 1925 had obtained void volumes of 30.6 and 36 percent, rodded and loose, respectively.[4–6] For wet sand the void volume (now occupied by water) rises to a maximum about 51 percent at 6 percent water and then slowly drops to the value for rodded dry sand at 17½ percent water. White and Walton illustrate a commercial sand (Wisconsin residual) which occurs in almost uniform spheres, though most commercial refractories are semiregular polyhedra. The void volume was found to be 37.7–42.5 for sand; packing improved with decreasing particle size. In general, jarring in vacuum improves packing perhaps by removal of air film. For silicon carbide, mullite, and alumina, they obtained a voidage of 40–50 percent, angular material giving a result a few parts percent higher. With mixtures of unlike materials they were successful in reducing this voidage, for lead shot with graded silicon carbide fines to 15.1 percent, in a rammed mix of silicon carbide and clay, as low as 10.7 percent. For porosities of 35.9 percent (tamped), 37.2 percent, 42.6 percent, 44.0 percent, and 44.7 percent in lead shot Foote, Smith, and Busang[4] in 1929 found corresponding coordination numbers 9.51, 9.14, 8.06, 7.34, 6.92, by an ingenious corrosion method. Horsfield[2] (1934), from an interesting study of the geometry of the spher-

FIG. 5. Particles of pure iron made by carbonyl decomposition and suitable for use in powder metallurgy. ×10,000.

ical aggregate, concluded that a mixture of six grades of spheres in correct proportions and stoichiometric arrangement could reduce voidage to 3.9 percent (Table II). How far practice diverges from computation is shown by the figure of 32.2 percent for an actual mixture of alumina particles, theoretically graded in accordance with these conclusions. This is not greatly different from the values obtained in rule-of-thumb mixtures, in alumina, and silicon carbide.

It is not difficult to see the cause of this discrepancy. Introduction of a smaller grade will destroy the regularity of the original aggregate, as soon as jarring begins. This will fair off the corners of the curve, and lower the over-all density. For example, introduction of an equal number of balls of radius $(\sqrt{2}-1)r$ or $0.4142r$ would certainly not give the wholly regular interstitial insertion postulated in the present calculation, no matter how long shaking lasted. Besides, introduction of an equal number of balls of radius $0.42r$ or $0.43r$ would presumably improve the density by some amount. Would this improvement extend to $0.5r$ or even further?

### Asymmetric Arrangements

So far we have ignored all patterns in the square holes except those with cubic symmetry, because of the laborious arithmetic involved. But it is likely that in the interval between the excellent packs of nine and twenty-one balls, (Fig. 4) some irregular or regular pattern would give a better arrangement than the *slack* degenerate $3^3$. The result would be to fill up the valleys on the graph with subsidiary smaller peaks. As $R/r$ decreases and $n$ increases, close packing within the interstices can be approached more closely. At the same time the disturbance

of the original close-packed array and *consequent decrease in effective density*, caused by a small intrusive particle, like the pea in the pilgrim's boot, will be enormously enhanced. The author considers that a smoothed curve, derived from Fig. 4, with faired corners, hills, and valleys, would be a most useful guide in operations like bulk storage, ceramic manufacture, and concrete mixing.

Figure 5, taken with the electron microscope, shows how perfect are the minute spheres of iron prepared by decomposition of carbonyl and actually used in powder metallurgy.

### Metallic Solid Solutions and Interstitial Compounds

Perhaps the nearest approaches to our model when $(R/r)=(\sqrt{2}-1)$, i.e., for one small ball per square hole, are interstitial solid solutions, like that of carbon in the high temperature $\gamma$-modification of iron, often called *austenite* or interstitial compounds like $Fe_3N$. This *austenite* persists down to 725°C and may be preserved to lower temperatures, at which it is metastable, by rapid cooling. If one-third of the iron atoms be replaced by nickel, the interstitial solid solution of carbon is stable down almost to room temperature.

### Hägg Rules for Interstitial Phases in Metallic Systems[10]

If the small ball is to be introduced without disturbing the original close-packed array, its radius must not exceed $0.4142r$. For the "heavy metals" hydrogen is the only element whose atomic radius would satisfy this requirement. From experimental study of a number of interstitial intermetallic phases, Hägg deduced an actual limit of about 0.59; obviously the lattice is capable of some adjustment. The rule holds very well for elements in the valleys of Lothar Mayer's curve, Sc-Ni, Y-Pd, Hf-Pr, which form with the "metalloid" elements H, B, C, N (with small atoms), interstitial phases based on the intermetallic compounds $M_4X$, $M_2X$, $MX$, $MX_2$. The second and third are most usual, the $M_2X$ phase has often a wide range of composition (like $\beta$-brass). $MX_2$ is formed only when the radius ratio is very small, with carbon, for example. In this phase the smaller atoms are often associated in pairs, as in $CaC_2$, which is tetragonal with $b/a$ equal to 1.6. Probably this pairing accounts for the elongation of the lattice to yield the elongated face-centered tetragonal structure, by a stretch in one direction. Of the two close-packed lattices, the cubic in general more readily forms MX phases, and the hexagonal $M_2X$. In general, these interstitial compounds are "metallic" in character; some—MoC, NbC, TaC, VN, TiN—even show supraconductivity.

The interstitial atoms arrange themselves so as to be near as many atoms as possible; they seem to strive after a coordination number of 6 (square hole), though 4 (triangular hole) is also known. It is probable that the intrusive atoms are crystallographically equivalent, though recent work (1946) by Jack[11] has shown that some order occurs in the iron nitride phase. A radius ratio exceeding Hägg's 0.59 does not predicate entire absence of interstitial phases; it means that those which are formed will have a much more complicated structure, e.g., carbides of Cr, Mn, Fe, Co, and Ni. In some properties these resemble ionic lattices.

In the author's view such complex interstitial compounds, usually with radius ratio exceeding 0.60, may quite reasonably be regarded as transitional structures. They lie between interstitial solid solutions, with random distribution like *austenite*, on the one hand, and the complex $\gamma$-brass structures formed between metals with radius ratio near unity, on the other. In turn, the $\gamma$-brass structure $(Cu_5Zn_8)_4$ links up through the ordered and random arrangements of the $\beta$-brass structure (based on CuZn) with the wide primary substitutional solid solutions formed by metals not very different in atomic radius, and of near-identical valency. Of these $\alpha$-brass and the continuous solid solutions in Cu-Ni alloys are striking examples.

Thus we can trace a gradation of properties:—Interstitial solid solution→interstitial compound→complex interstitial structure→$\gamma$-brass structure→ordered $\beta$-structure→random $\beta$-structure→primary substitutional $\alpha$.

[10] G. Hägg, Zeits. f. physik. Chemie **12**, 33 (1930–31).

[11] K. H. Jack, Nature **158**, 60 (1946).

# Resistance of Oxide Cathode Coatings for High Values of Pulsed Emission*

W. E. Danforth and D. L. Goldwater
*Bartol Research Foundation of the Franklin Institute, Swarthmore, Pennsylvania*
(Received July 27, 1948)

Measurements are described of the potential variation of fine ribbon probes embedded in standard barium and strontium oxide coatings. Observations were made with 19 tubes at values of pulsed emission as high as 50 amperes/cm². Results were as follows: (a) values of resistance were found ranging from 2 to 230 ohms-cm²; (b) potential gradients were adequate to admit dielectric breakdown as a cause of sparking; (c) the curve of resistance *vs.* emission passes through a maximum and, at different temperatures, this maximum occurs at the same probe potential; (d) with BaO or mixed cathodes of higher values of resistance the resistance is largely at the interface, for low resistance coatings no observations on this matter were obtained; (e) the two best tubes as regards high emission without sparking had the lowest values of coating resistance; (f) the resistance of SrO cathodes is an order of magnitude higher than that of BaO or mixed cathodes; (g) superposition of d.c. upon pulsed emission gives a pronounced decrease of resistance.

## 1. INTRODUCTION

THIS work was performed in the winter of 1944–45 as part of an investigation concerning the mechanism of sparking in magnetrons using oxide coated cathodes. An hypothesis was being considered that the predominating cause of sparking, at least in diodes, was a disruptive dielectric breakdown of the oxide coating due to a potential difference resulting from the finite resistance of the coating material or of the coating-to-base-metal interface. Subsequent work[1] showed that the anode plays a somewhat more prominent role than was thought at first, and that a spark may be either anode-initiated or cathode-initiated although in both cases the most brilliant visual manifestations occur at the surface of the cathode. Even when a spark is anode-initiated, however, the cathode coating, by virtue of its resistance, plays a supporting role and the reduction of such resistance is considered advantageous.†

Resistance measurements were effected by means of embedded probes following the method of Becker and Sears.[2] Shortly before pressure of other matters caused us to discontinue this investigation the subject was taken up by the Cathode Research Group at the M.I.T. Radiation Laboratory, with whom we were in collaborative relationship. The results of their work have already been published.[3] Other publications on this subject include that of Wright[4] who gave careful attention to boundary layer effects at relatively low currents, and Loosjes and Vink[5] who investigated cathode drop by means of a movable anode.

## 2. TECHNIQUES

### A. Construction of the Cathodes

Cathodes were constructed according to the diagram of Fig. 1. The weighed nickel sleeve with thoria-embedded heater received an oxide coating, then several turns of nickel ribbon ($\frac{1}{2}\times 5$ mil cross section), another coating, a second probe, and a final coating. Appropriate weighings during fabrication enabled weights of both coatings to be determined. To prevent the probes from uncoiling, when the carbonate was converted to the oxide,

* This paper is a condensation of O.S.R.D. report NDRC 14–514 Cathode Coating Resistance as Measured by Embedded Probes, W. E. Danforth, Oct. 31, 1945.

[1] W. E. Ramsey, "A General Survey of Sparking Phenomena in High Vacuum Thermionic Tubes," O.S.R.D. report, Contract OEMsr-358, Bartol Research Foundation, October 1945.

† The question may be raised as to whether the measurements here reported really give quantities which should be termed "resistance." In this connection we may state that we here define *resistance* as that quantity by which the current density through the coating shall be multiplied in order to give the observed probe voltages. Since the oscilloscope used for voltage measurements was of impedance very high compared to that between probe and base metal, the authors feel justified in neglecting any errors which might be caused by current carried by the probe. Also it was determined that e.m.f.'s of thermal or electrochemical origin between probe and base metal were very small compared to the observed voltages which resulted from the drawing of current through the coating.

[2] J. A. Becker and R. W. Sears, Phys. Rev. **38**, 2193 (1931).

[3] A. Fineman and A. S. Eisenstein, J. App. Phys. **17**, 663 (1946).

[4] D. A. Wright, Proc. Roy. Soc. **190**, 394 (1947).

[5] R. Loosjes and H. J. Vink, Philips Res. Rep. **2** 190–204 (1947).

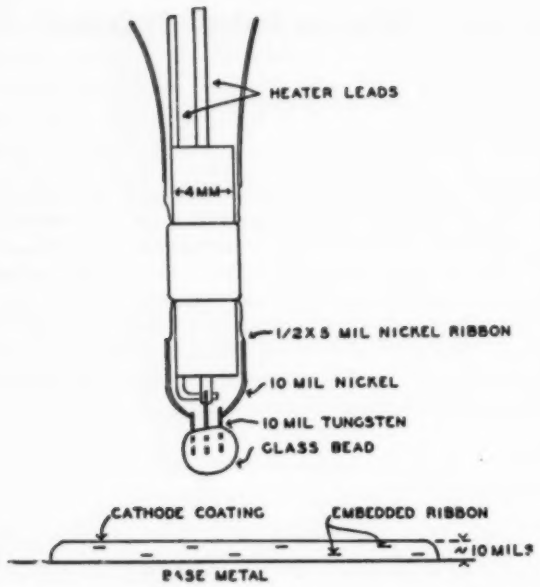



FIG. 1. Structure of cathode with two helical nickel probes embedded in coating.

they were secured at both ends by spot welding to nickel supports. The sleeve material was Grade *A* nickel, and the coated area was 0.63 sq. cm.

As regards the weight of coating which one applies, it became evident that at least 5 mg/cm² were required between the inner probe and the base metal if short-circuiting was to be avoided and that even 10 mg/cm² were desirable. The weight to be applied between the two probes should be of the same order as the first coat. The third coat, which assists in holding the second probe in place, must also be of the same order if mechanical troubles are to be avoided. It was therefore necessary that a double-probe cathode have a total weight of approximately 20 mg/cm², an amount of coating approximately double that which we were accustomed to apply to an oxide cathode for high pulsed emission.

### B. Tube Preparation

For purposes of this study the cathodes were mounted on four lead presses in glass bulbs. The anode was a flanged Kovar cup of 11.7-mm I.D. and 1-inch long and could be immersed in running water during tests. All tubes used in this study were processed on a system which employed a mercury diffusion pump. The tubes were first made up with a tungsten filament mounted in place of the probe-cathode by means of which the anodes were subjected to electron bombardment before the oxide cathodes were installed. The anodes were maintained at a dull red heat for about one-half hour, after which time the bombardment produced no appreciable increase of pressure. The tube was then cracked open and the probe cathode installed. The process of outgassing and activation was as follows:

After sealing-on, the system was pumped to a pressure less than $10^{-5}$ mm when an oven was raised so as to envelope the experimental tube, an ion gauge, and a liquid air trap which separated the tube and the ion gauge, from the rest of the system. (In addition to this "second" trap there was, of course, a primary trap on which liquid air was kept constantly as long as a tube was on the system.) The temperature was raised to 450°C in about half an hour, maintained there for one hour, and allowed to cool for perhaps another hour. The pressure, as measured on an ion gauge located between the second liquid air trap and the primary trap, would rise to some value between $10^{-4}$ mm and $10^{-3}$ mm during bake out and fall back to $1\times10^{-7}$ mm after the system had cooled.

Following the bake, a small heater current was applied to the cathode, so as to raise its temperature to about 650°C at which temperature it would be allowed to stand overnight for outgassing of the cathode parts. The next morning further outgassing was accomplished by heating the lower two-thirds of the cup to redness by means of a resistance heater designed to slip over the cup. The cup was kept at a temperature of about 800° for one hour. After this process the tube was regarded as ready for activation.

In the majority of the tubes (all, in fact, except those where conversion was made the subject of probe measurements) conversion of the carbonates to oxides and activation of the oxides was performed at approximately 1000°C brightness temperature. The procedure was to set the pyrometer for 1000° and raise the cathode temperature as rapidly as possible to this value. Because of the rather large heat capacity of our cathode design it would require a minute or two to do this. In general the temperature would be allowed to remain at 1000° for about 30 seconds

and then reduced to 950°. If, upon again raising the temperature to 1000°, the pressure did not increase to more than about 2 or $3\times10^{-7}$ mm, it was judged that conversion was complete. If, however, further increase of pressure occurred, the cathode would be allowed to remain at 1000° for another period of approximately 30 seconds and then be reduced to 950°. In most cases repetition of the 1000° treatment was unnecessary.

The above purely thermal treatment resulted in a considerable degree of activation of the cathode for, keeping the temperature at 950°, it was found that 50 ma could be drawn, space charge limited, with no subsequent decay of emission. Before removing the tube from the vacuum system it was subjected to a period of d.c. drain of at least 50 ma for at least one hour. In many cases the treatment included a 50-ma treatment, followed by a 90-ma treatment followed by a 120-ma treatment, each for at least one-half hour.

As would be expected by anyone familiar with oxide cathode activation, the above mentioned d.c. treatment was accompanied by a release of gas, each increase of the current resulting in a new burst of gas. That some of this gas came from the anode, which had, of course, been exposed to air after its preliminary outgassing, was often demonstrated by a fall of pressure when the draft from a blower was directed at the Kovar cup. Toward the end of this investigation observations were made of cathode resistance during this activation process and a discussion of these matters will appear later.

Following the activation, it was our practice with the earlier tubes to give them another bake to the extent of raising the temperature to 300° and then allowing the system to cool, the whole process taking about one hour. This was abandoned for the later tubes and no change of performance was noticed.

Before sealing the tube off, liquid air was placed on the second trap (a trap which was very close to the tube and subjected to the same baking) which caused the pressure, under best conditions to fall as low as $2\times10^{-8}$ mm. (The pressure here referred to was measured on an ion gauge directly adjacent to the tube.)

### C. Probe Measurements with 1-Microsecond Pulses

Inasmuch as the problem of sparking in pulsed magnetrons was one of the major concerns of the present investigation it was obvious that the resistance measurements should be made with microsecond pulses. Figure 2 shows the scheme of connections as made to a modulator (MIT Radiation Laboratory Model 4; 27,000-volt peak). Because of the necessity for keeping the cathode at ground potential, and since the sign of the high voltage output of the modulator is negative, a pulse transformer provided by the Radiation Laboratory was used to permit grounding the negative side. The high positive terminal goes directly to the anode of the tube (Kovar cup) and the cathode sleeve is grounded. The pulsating probe potentials are observed by connecting the plates of a synchroscope (No. 2) directly between the probe and the cathode sleeve. Another synchroscope (No. 1) served to measure the anode voltage by means of a resistance divider. Simultaneous observations of anode voltage and emission current were not made, only one synchroscope being available for both of these observations. Included in the diagram is a 400-volt power supply which was used in observing the effect upon probe behavior of the superposition of a direct current. A high capacity across its output insured a low impedance bypass for the pulsed currents.

Table I gives coating weights and maximum space charge limited emission for all tubes which were successfully brought to a point where measurements of one kind or another could be made. The meaning of the symbols at the heads of columns is as follows:

$m_1$ = coating weight between sleeve and inner probe,
$m_2$ = coating weight between inner and outer probes,

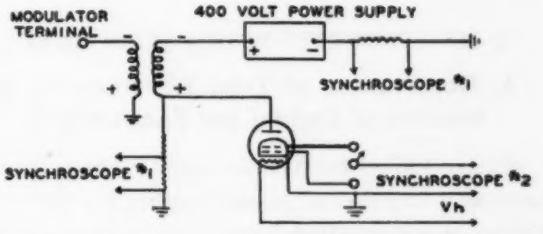



FIG. 2. Circuit arrangement for observation of probe voltages while drawing pulsed emission from cathode. The 400 v power supply permits superposition of d.c.

TABLE I.

| 1 Tube No. | 2 $m_1$ Mg/cm² | 3 $m_2$ Mg/cm² | 4 $m_3$ Mg/cm² | 5 $I_m$ amp./cm² |
|---|---|---|---|---|
| *P* 1 | 20 | — | 20 | 20 |
| *P* 2 | 4.7 | 12.7 | — | 20 |
| *P* 3 | 4.7 | 14.3 | — | 26 |
| *P* 4 | 3.0 | 7.6 | 12.1 | 27 |
| *P* 6 | 7.1 | 3.2 | — | 40 |
| *P* 7-2 | 7.1 | 4.5 | 6.2 | 53 |
| *P* 8-2 | 5.4 | 3.5 | 7.0 | 21 |
| *P* 10-2 | 8.6 | 4.0 | 2.8 | — |
| *P* 11 | 5.1 | 6.8 | 7.5 | 17 |
| *P* 13 | 5.7 | 5.2 | 11.0 | 10 |
| *P* 15 | 5.7 | 6.7 | 6.5 | 10 |
| *P* 16 | 5.9 | 6.5 | 12.1 | 25 |
| *P* 17 | 8.4 | 8.1 | 11.4 | 50 |
| *P* 19 | 9.4 | 7.0 | 17.5 | 12 |
| *P* 20 | 13.7 | 1.1 | 20.8 | 9.6 |
| *P* 21 | 16.7 | 4.1 | 16.0 | 9.7 |
| *P* 24 | 20.3 | — | 16.5 | 18 |
| *P* 25 | 10.8 | — | 13.3 | 19 |
| *P* 27 | 18.1 | — | 17.9 | 15 |

$m_3$ = coating weight over the outer probe,
$I$ = maximum space-charge-limited emission at 850° as obtained from a $V-I$ plot on two-thirds-power paper.

It is seen that, as regards the emitting power of these tubes, they may be said to be of medium grade, somewhat inferior to the best laboratory-activated cathodes and better than most production cathodes. The fact that their emission is less than our best diodes is probably due to some circumstance associated with the probe. This was shown by tests in which cathodes, exactly similar to the probe cathodes but without probes, were activated with the same procedure and the same equipment as the probe cathodes, and gave emissions exceeding 50 amps./cm².

The missing numbers in the above list correspond to tubes which for some reason had to be accounted for as failures. The dashed numbers (e.g. *P* 7-2) indicate tubes which failed initially and were re-processed with a new cathode.

## 3. MEASUREMENTS WITH 1 μSEC. PULSES

### A. Observations of Total Resistance as a Function of Current and Temperature

Figure 3 gives for 8 tubes, total cathode resistance as a function of pulsed current for different values of temperature.

All tubes here described excepting *P* 1 had double-probe cathodes and, unless otherwise noted, used equimolar barium and strontium carbonates.

Since the double-probe technique purports to discriminate between resistance of the coating itself and resistance of the interface, some appropriate definitions should be adopted.

In this section the resistance of the coating itself will be referred to as *coating resistance*, and that of the boundary as *interface resistance*. Their sum will be called *total resistance* or *cathode resistance*.

The figure which we have referred to gives curves of total resistance as measured with the outer of the two probes. Values of coating weight between the outer probe and the sleeve are given on each figure. It will be noticed that all tubes show that as the current increases the resistance first increases, then passes through a maximum and decreases. The value and position of the maximum shows considerable variation from one tube to another. Table II shows, for 11 tubes, all at approximately 850°C, values of maximum resistance and values of current at which the maximum occurs. It is seen that the maximum resistance has values between 3.0 and 100 ohm-cm² (a factor of 33.3) and that the position of the maximum lies between 2.4 and 31 amp./cm² (a factor of 13).

The column at the extreme right gives values of the probe voltage at which the maximum occurs and it is noticed that this figure shows a considerably greater constancy than either of the other two, its value lying between 75 and 240 volts (a factor of 3.2).

Further evidence that the maximum tends to come at a definite value of probe voltage is given in Fig. 4, which shows the same data as Fig. 3, except that the resistance is plotted against voltage on the probe instead of emission current. One notices that, when plotted in this manner the picture is considerably simplified, the maxima at different temperatures falling one above the other.

### B. Results of Double-Probe Measurements

Figure 5 shows a plot of data from four of the tubes in which both probes showed consistent behavior.

Assuming that the coating resistance is constant throughout the material and that the re-

sistance of a layer of coating is proportional to its weight, one may calculate the interface resistance as follows:

$$R_0 = R_1 - (R_2 - R_1)(m_1/m_2),$$

where

$R_0$ = interface resistance,
$R_1$ = total resistance to inner probe,
$R_2$ = total resistance to outer probe,
$m_1$ = weight of coating between the inner probe and the base metal,
$m_2$ = weight of coating between probes.

Referring again to the figures, the dotted curves denote measurements on the inner probe and the solid curves, the outer probe. Using the above formula we proceed to calculate the interface resistance for these four cases. The results are shown in Table III, in which column 1 gives the tube number, column 2 the value of $m_1/m_2$, column 3 the value of current at which the calculation is made. Column 4 gives the values of the total resistance shown by the inner probe; column 5 gives the resistance between probes; and column 6 gives the calculated interface resistance. The values of current given were chosen to correspond to the maximum resistance except for the case of $P$ 11 where data do not include the maximum and 10 amp./cm² was arbitrarily chosen. The data at approximately 850° were used in all four cases, and all cathodes

TABLE II.

| Tube No. | Value of maximum ohms-cm² | Position of maximum amp./cm² | Probe voltage at maximum |
|---|---|---|---|
| $P$ 1 | 46 | 3.5 | 161 |
| $P$ 2 | — | — | — |
| $P$ 3 | — | — | — |
| $P$ 4 | — | — | — |
| $P$ 6 | 14 | 15 | 210 |
| $P$ 7-2 | 3.0 | 25 | 75 |
| $P$ 10-2 | 10 | 18 | 180 |
| $P$ 13 | — | — | — |
| $P$ 15 | 27 | 8.0 | 216 |
| $P$ 16 | 16.2 | 5.5 | 89 |
| $P$ 17 | 3.5 | 31 | 108 |
| $P$ 19 | 7.3 | 23 | 168 |
| $P$ 24 | 30 | 5.4 | 162 |
| $P$ 25 | 38 | 5.0 | 190 |
| $P$ 27 | 100 | 2.4 | 240 |
| | | | Average 163 |

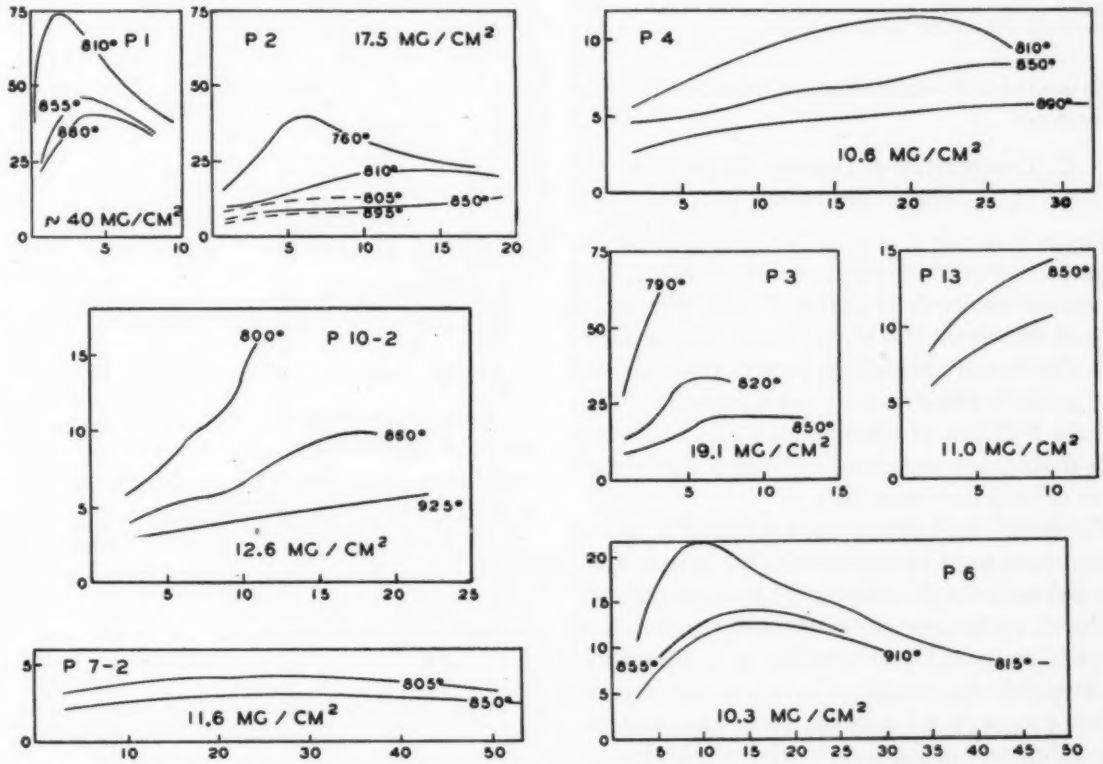



FIG. 3. Resistance of cathode coating (including interface) as function of pulsed emission current. Ordinates are ohms-cm²; abscissae, amps./cm².

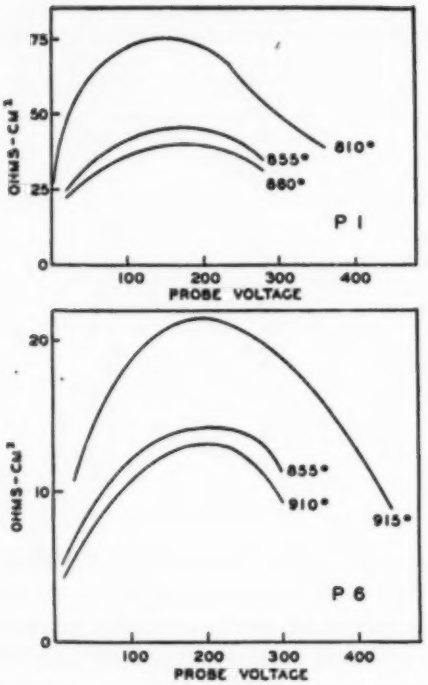



FIG. 4. Resistance of cathode coating (including interface). In these graphs resistance is plotted against the voltage which appears at the probe as the emission current is varied. Range of variation of emission current is the same as shown in Fig. 3 for the same tubes.

were coated with equimolar barium and strontium carbonates.

### C. Comparison of Barium Oxide with Strontium Oxide

Figure 6 shows a comparison between the resistance of barium oxide cathodes with that of strontium oxide. It is apparent that the resistance of the latter is a higher order of magnitude than the former. Materials in both cases were of C.P. grade without further purification.

Tube *P* 19 had a barium oxide cathode in which both probes were working and hence can yield a value of interface resistance.

The strontium oxide cathode in tube *P* 20 shows a maximum total resistance of 230 ohms at 890°. The values given by the two probes are indistinguishable within our experimental uncertainty of about 10 percent. Unfortunately this cathode was one in which the amount of material between the probes was only 8.1 percent of the amount between the inner probe and the sleeve. This fact, in connection with the above-mentioned experimental uncertainty, makes it impossible to draw conclusions regarding the relative importance of coating resistance and interface resistance.

### D. Effect of d.c. Component

Figure 7 shows the effect upon total resistance of the superposition of a d.c. component upon the pulsed current. The data for *P* 1, *P* 2, and *P* 7-2 show resistance plotted against pulsed current for different values of superimposed d.c. In the case of *P* 3 the pulsed current was held constant at 10 amp./cm² and the variation of resistance with d.c. component was recorded. Values of d.c. component have not been reduced to current per unit area but refer to the standard area of 0.63 sq. cm.

### E. Results with Two Probes at the Same Level

Figure 8 shows results obtained with tubes which were made up for the purpose of testing the reliability of our probe measurements. The cathodes were made in the identical manner as those in the other tubes except that both probes were at

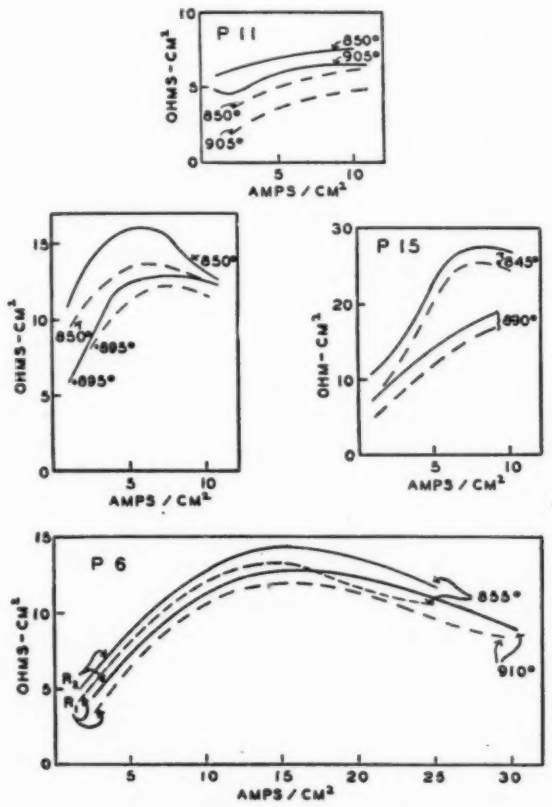



FIG. 5. Resistance of cathode coating (including interface) in 4 cases where successful runs were made on both probes.

the same depth in the coating and hence should both give identical curves. The probes were wound on as a double thread, the convolutions of one of the helices lying between those of the other, so that no possibility of accidental contact between the probes existed. It is seen that differences up to 10 percent are experienced. The values of coating weight lying between the base metal and the probes are given on the figures.

### F. Time Changes of Resistance Observed with 1 Microsecond Pulses

Observations of the probe voltage pulses, by means of the arrangement in Fig. 2 often show distortion of the square wave to such a form as would be produced by an increase of resistance with time during the microsecond. Figure 9 shows the type of pulse which characterizes this phenomenon. In general this rise of probe voltage with time, for constant current through the cathode is a phenomenon which tends to appear at high current densities. For example, a tube would be likely to give square probe pulses at 5 amps./cm² and pulses of the above shape at 20 or 30 amps./cm². It was noticed in some cases that this effect developed as the tube aged. The occurrence of sparking in the tube was also likely to bring about this effect in a tube originally free from it.

In many respects the onset of this phenomenon reminds one of the onset of decay in a cathode originally free from it. Time did not permit an investigation of any possible correspondence between these two effects.

### G. Measurements with Prolonged Pulses

During the course of this work, a series of experiments was carried out by W. E. Ramsey concerning variations of resistance with time when the cathodes were subjected to pulses of several tens of microseconds duration.

TABLE III.

| 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| Tube No. | $\frac{m_1}{m_2}$ | $I_{max}$ amp./cm² | $R_1$ | $R_2-R_1$ | $R_{interface}$ ohm-cm² |
| $P$ 6 | 2.25 | 15 | 13.3 | 1.3 | 10.4 |
| $P$ 11 | 0.74 | 10 | 6.3 | 1.3 | 5.3 |
| $P$ 15 | 0.86 | 8 | 25.4 | 2.2 | 23.5 |
| $P$ 16 | 0.90 | 5.5 | 13.6 | 2.6 | 11.3 |

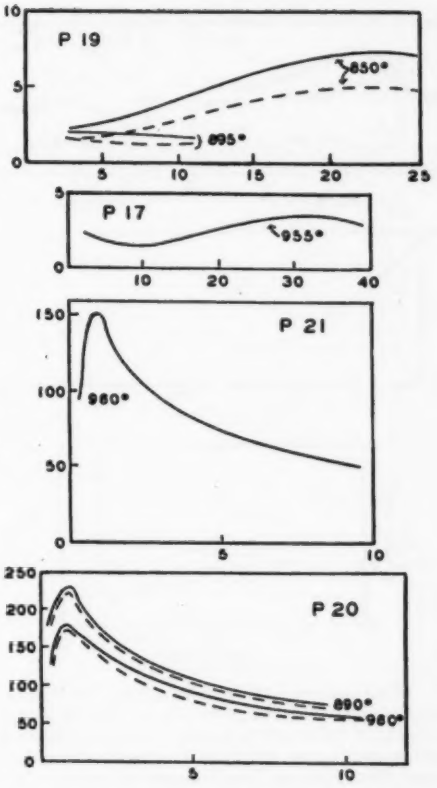



FIG. 6. Resistance of cathode coating (including interface). Comparison of BaO coating ($P$ 17 and $P$ 19) with SrO coating ($P$ 20 and $P$ 21). Note the much higher resistance of SrO.

A considerable complexity of phenomena was observed, which will be presented in a future paper.

In general, it may be said that the resistance increases with time as high pulsed current is applied, in agreement with the above-mentioned findings with 1 microsecond pulses. It was further observed, however, that superposition of a small d.c. had the effect of preventing the increase with time.

### H. Correlation of Sparking with Resistance

It was largely to ascertain whether or not the above-named correlation existed that this research was undertaken. The type of sparking here concerned is that which takes place with 1 microsecond pulses in the space-charge-limited state. Table IV gives the status of this question at the time when the work was suspended in favor of other matters.

The meanings of the eight columns are as

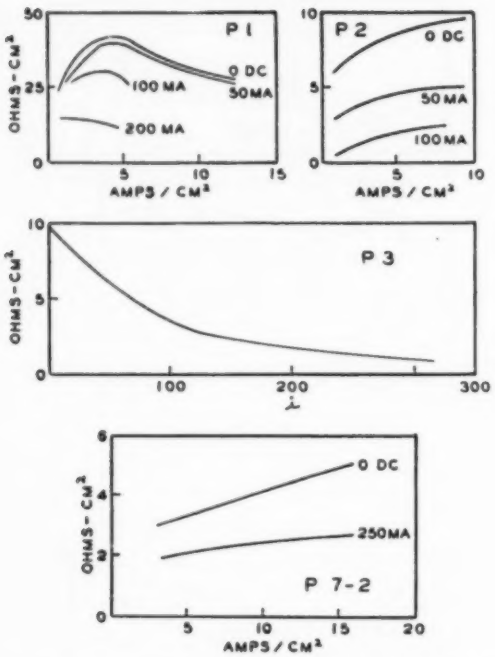



FIG. 7. Resistance of cathode coating (including interface). Results illustrate the pronounced decrease of coating resistance which takes place when d.c. is superposed upon the pulsed emission. In the case of $P$ 3 abscissae represent d.c. current from 0.63 sq. cm of cathode surface.

follows: (1) Tube number, (2) Centigrade brightness temperature, (3) Pulsed current at the onset of sparking, (4) Probe voltage (outer probe) at sparking, (5) Power dissipated in the coating between core and outer probe, (6) Resistance to the outer probe at sparking, (7) Weight of coating under the outer probe, (8) Constitution of the coating material. The symbol Ba signifies a barium oxide coating; Sr, strontium oxide; and Ba-Sr, an equimolar mixture.

As regards the 8 Ba-Sr tubes up to and including $P$ 16, one sees that the one tube which gave an unusually high sparking current ($P$ 7-2, 71.5 amps./cm²) also had an unusually low resistance (2.8 ohms-cm²). The remainder of the Ba-Sr tubes, sparked between 29 and 43 amps./cm², had resistances between 8.06 and 15.5 ohms-cm², and showed no correlation. Tube $P$ 17, a barium oxide tube, had the second highest emission (50 amps./cm²) and also a rather low resistance (5.08 ohm-cm²).

The correlation between sparking current and resistance is pronounced in the case where one considers Sr on the one hand and Ba-Sr or Ba on the other. In the former case, the sparking currents are 9.5 and 9.7 amps./cm² and the resistances 69 and 49 ohms-cm², while in the latter case the sparking currents lie between 25 and 71 amps./cm² with the resistances between 8.06 and 15.5 ohms-cm². The potential difference (probe voltage, $V_p$) which the coating can support without sparking appears to be somewhat higher in the case of Sr.

Regarding the particular mechanism through which high resistance favors the occurrence of sparking, we have been inclined to think of it as dielectric breakdown of the coating or interface. Positive ions ejected in this event could form local perturbations of the space charge sheath and give rise to a localized instability which would result in an arc.

The simplest picture based on the above theory would require that, in a number of tubes having different coating resistance (operated at the same temperature) sparking would occur in each case at that current which would produce a certain $IR$ drop in the cathode layer. It will be observed from Table IV that values of $IR$ (i.e., $V_p$) at sparking vary by more than a factor of 2 and that this spread is not correlated with differences in coating weights. However, differences in dielectric strength from tube to tube are indeed to be expected.

A. S. Eisenstein[6] has stated that sparking tends to occur at a definite value of $I^2R$, rather than of $IR$, suggesting that the mechanism involves a thermal instability rather than dielectric breakdown. We find ourselves reluctant to abandon the

TABLE IV.

| $P$ | $T$ | $I_s$ amps./cm² | $V_p$ | $I_sV_p$ peak watts | $R_s$ ohms-cm² | $W$ Mg/cm² | |
|---|---|---|---|---|---|---|---|
| 3 | 848 | 31.8 | 286 | 9,100 | 9.00 | 19.1 | Ba-Sr |
| 6 | 855 | 31.4 | 305 | 9,580 | 9.72 | 10.3 | Ba-Sr |
| 7-2 | 850 | 71.5 | 200 | 14,280 | 2.80 | 11.6 | Ba-Sr |
| 10-2 | 848 | 34.9 | 282 | 9,840 | 8.06 | 12.5 | Ba-Sr |
| 11 | 843 | 30.5 | 263 | 8,020 | 8.53 | 11.9 | Ba-Sr |
| 13 | 847 | 43.0 | 362 | 15,580 | 8.4 | 11.0 | Ba-Sr |
| 15 | 847 | 33.3 | 517 | 17,220 | 15.5 | 12.4 | Ba-Sr |
| 16 | 857 | 29.6 | 277 | 8,200 | 9.35 | 12.4 | Ba-Sr |
| 17 | 847 | 50.0 | 254 | 12,700 | 5.08 | 16.5 | Ba |
| 19 | 850 | 25.0 | 180 | 4,500 | 7.20 | 16.4 | Ba |
| 20 | 890 | 9.5 | 660 | 6,260 | 69 | 14.8 | Sr |
| 21 | 960 | 9.9 | 474 | 4,600 | 49 | 20.8 | Sr |

[6] A. S. Eisenstein, M.I.T. Radiation Laboratory Report 933.

latter hypothesis, however, for the following reasons:

(a) Values of $V_p$ at sparking give reasonable breakdown field strengths for a "poor" dielectric.

(b) The sparking here concerned can take place in a fraction of a microsecond. We question whether the thermal hypothesis can account for these short times. The values of $I^2R$ in the present data are all less than 20,000 watts. Therefore the maximum heat which can develop in a coating in one microsecond is $5\times10^{-3}$ calories per square cm.

(c) Constancy of $I^2R$ from tube to tube at sparking does not necessarily disprove the breakdown hypothesis. If the cathode coatings which yield higher values of current at sparking also happen to have lower values dielectric strength, a fortuitous constancy of $I^2R$ could conceivably result.

### I. Variation of Resistance with Temperature

No extensive quantitative study of the effect of temperature upon resistance has as yet been made in this investigation. It is obvious by inspection of Fig. 3 that a considerable variety of phenomena are encountered.

Theoretical considerations regarding impurity semi-conductors, however, suggest that the conduction should follow a relation of the form $e^{-Q/2KT}$, provided of course that the constitution of the material does not itself change with temperature.

Figure 10 shows the data for tube $P$ 2 at 6 amperes pulsed current plotted for 3 temperatures. Plotting the logarithm of $1/R$ against the reciprocal of absolute temperature one has, according to the above formula, a straight line whose slope is proportional to the excitation energy $Q$. Line A corresponds to resistance between the probes while B refers to resistance between the inner probe and the sleeve. The values of Q given on the figure are within the spread of data obtained by others.[7]

### 4. MEASUREMENTS DURING CONVERSION AND ACTIVATION

Observations have been made of the changes of resistance which take place during activation of the cathodes. Some general comments are as follows:

(a) It is evident that conversion by heat below 800°C of the barium or strontium carbonates to oxides does not immediately result in a conducting or emitting substance.

[7] J. P. Blewett, J. App. Phys. 10, 668 (1939).

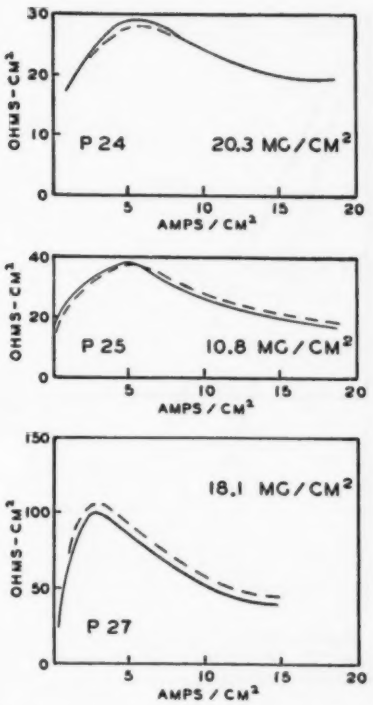



FIG. 8. Resistance of cathode coating (including interface). These curves were obtained with two probes at same level and illustrate the order of accuracy associated with the foregoing observations.

(b) Holding the oxide in the neighborhood of 800°C for a sufficiently long time will, however, cause a fall of resistance. In all cases examined, the appearance of a certain amount of emission accompanied the fall of resistance.

(c) Although a cathode may be activated in a short time by thermal treatment alone at 900° to 1000° it appears that further decrease of resistance can always be affected by the drawing of electron current.

(d) During activation of a cathode by application of d.c. in successively increasing values it is often found that each increase results in a fall of resistance to a new low value which, from then on, is likely to rise slowly.

Referring now to the specific experiments, consider Fig. 11 which describes the processing of

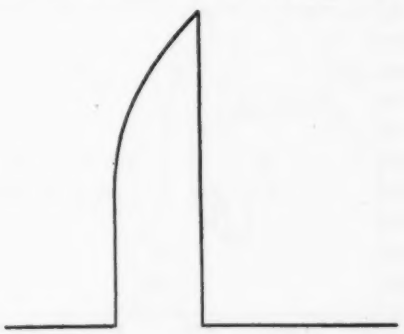


FIG. 9.

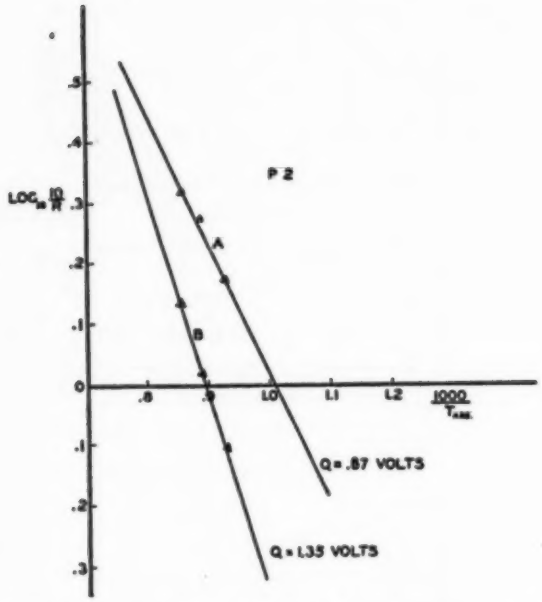



FIG. 10. Resistance as function of temperature. A, between probes; B, between inner probe and base metal. Cathode emission 6 amps./cm².

tube *P* 19, a barium carbonate coating. The figure comprises three graphs in which the several pertinent variables are plotted against the time during the manipulation of the tube while still on the vacuum system. The uppermost of the three graphs shows the specific manner in which the temperature of the cathode was increased during the process. Temperatures down to about 750° were measured by pyrometer. The values of lower temperature are only approximate, being obtained by plotting the curve of pyrometer temperature *vs.* heater current and making a straight-line extrapolation between the lowest point and the origin. The middle graph shows the rise of pressure, measured in an ion gauge in a position adjacent to the experimental tube, which occurred by reason of breakdown of the carbonate. All data are here plotted in terms of microamperes of positive ion current, the calibration figure being approximately $1\times10^{-5}$ mm of Hg per microampere.

It will be observed here that the breakdown, which was taking place only very slowly at 660°, became very much faster at 710° and went rapidly to completion when the temperature was increased to 740°.

The bottom graph in Fig. 11 shows at what point in the process the coating resistance fell and the emitting power developed. The procedure in this case was to apply a potential to the tube which would correspond to a space-charge-limited current of 10 ma and to observe the rise of current from zero to this value.

Resistance values, where emission of a few milliamperes or more was available, were measured by drawing current to the anode and measuring the probe potentials by a high resistance voltmeter.

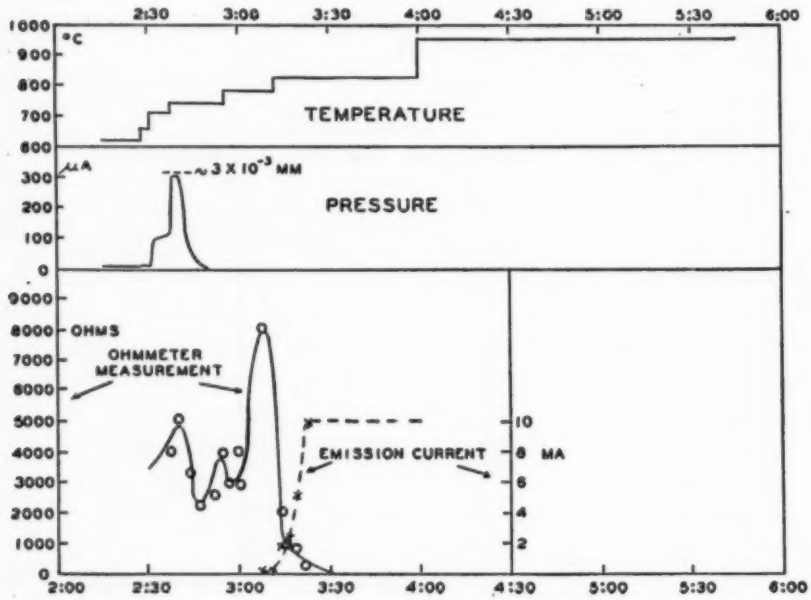



FIG. 11. Resistance observations during activation.

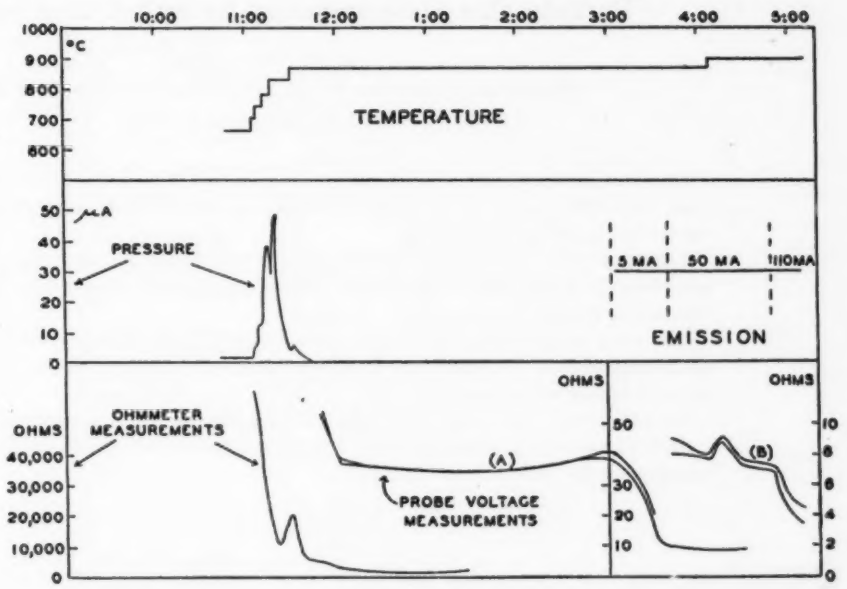



FIG. 12. Resistance observations during activation.

The curve marked "ohmmeter measurement" should not be considered in a strictly quantitative sense but only as an indication as to the point in the process where the resistance fell. Before activation started resistance values were obtained by applying a potential of 4.5 volts between the probe and the sleeve, with a series resistance of 4500 ohms, and noting the current. Care was taken always to use the same polarity since the resistance differed, in amounts approaching a factor of two, between the two directions.

Figure 12 was made with two probes, both situated at the same depth in the coating. The coating was equimolar barium-strontium carbonate. In this case no continuous current was drawn during the first stages of activation. The temperature was held constant at 870° for over four hours, and the resistance was observed to reach an equilibrium value in about one hour.

This case shows clearly that even after thermal activation at 870° has gone as far as it will, the resistance may be further reduced by a large factor by drawing current. After 3½ hours of treatment at 870° the coating resistance was 40 ohms. Drawing 5 ma reduced this to 9 ohms in ½ hour. In this case raising the current to 50 ma did not appreciably change the resistance nor did an increase of temperature to 900° have much immediate effect. When the current was increased to 110 ma, however, a pronounced fall of resistance occurred.

## 5. ACKNOWLEDGMENTS

It is a pleasure to acknowledge the helpful support of Dr. W. F. G. Swann, who was Official Investigator of the project in which this work was carried out.

The writers are indebted to W. E. Ramsey, M. A. Pomerantz, and C. D. Prater for helpful suggestions and assistance.

Construction of the probe cathodes was ably carried out by Henrietta W. Fricke. Matters connected with the construction and pre-outgassing of the tubes were taken care of by Austin G. Nester and J. Garrett Blauvelt.

# Particle Size Determination by Soft X-Ray Scattering*

K. L. YUDOWITCH**
*Department of Physics, University of Missouri, Columbia, Missouri*
(Received August 2, 1948)

Small angle x-ray scattering measurements are made on samples of colloidal gold of radius from 50 to 400 angstroms. Copper (1.54 angstrom wave-length) and aluminum (8.32 angstrom wave-length) radiation are used.

The usual method of analysis is shown to be valid only for particles of radius less than 120 wave-lengths. Extension of the method to larger particles is achieved, giving improved electron microscope correlation.

The use of longer wave-lengths and optimum shaped slits are shown to reduce geometry errors sufficiently to give clear evidence of predicted intensity maxima. The first three secondary shape function maxima are observed.

An additional interference maximum attributable to particle-to-particle interference is observed. The position and magnitude of this maximum are correctly predicted by an expression based on earlier fluid scattering theory.

SMALL angle x-ray scattering has been employed since 1939 as a means for determining size parameters of particles of dimensions less than 500 angstroms. Most work of this type is based on a theory such as that developed by Guinier[1–3] or Hosemann.[4,5] Certain assumptions are inherent in the application of this theory. Guinier first derived an expression for the scattering of a perfectly collimated monochromatic beam from a single particle having negligible absorption. By comparing Guinier's result with the similar expression obtained for the optical case of an opaque disk[6] (total absorption), the absorption effect may be seen to be truly negligible. With proper experimental technique, the effect of imperfect monochromatization is made negligible. However the collimation error is usually appreciable. A simplified comprehensive treatment of the collimation error is described in this paper.

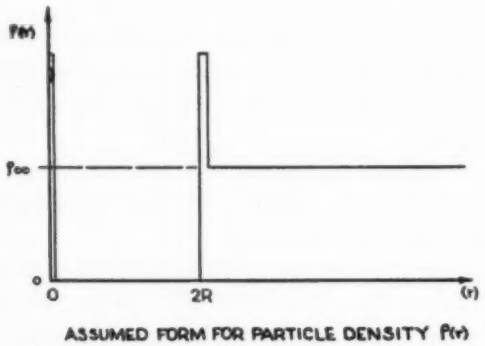



FIG. 1. Assumed distribution of particle density $\rho(r)$.

The next step in Guinier's derivation is the simple addition of the intensity scattered by each particle in the sample. This assumes multiple scattering, multiple refraction,[7] and interference between particles to be negligible. Most workers[8,9] agree that multiple scattering and multiple refraction are unobserved experimentally. Neglecting interference between particles however will be shown to be unjustified, at least in certain cases.

Non-uniformity in the size of most specimens gives rise to a condition not included in Guinier's simple theory. A number of workers[8,10–12] have devised methods for determining the size distribution from the experimental small angle scat-

* Extract from a thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the University of Missouri.

** Present address: Department of Physics, Florida State University, Tallahassee, Florida.

[1] A. Guinier, Annal. de Physique **12**, 161 (1939).
[2] A. Guinier, *Radiocrystallographie* (Dunod, Paris 1945), (Chap. XII).
[3] A. Guinier, J. de Chimie Physique **40**, 133 (1943).
[4] R. Hosemann, Zeits. f. Physik **113**, 751 (1939).
[5] R. Hosemann, Zeits. f. Physik **114**, 133 (1939).
[6] J. Slater and N. H. Frank, *Introduction to Theoretical Physics* (McGraw-Hill Book Company, Inc., New York, 1933), p. 324.
[7] R. Von Nardroff, Phys. Rev. **28**, 240 (1926).
[8] C. G. Shull and L. C. Roess, J. App. Phys. **18**, 295 (1947).
[9] W. W. Beeman and P. Kaesberg, Phys. Rev. **72**, 512 (1947).
[10] A. Guinier and G. Fournet, J. de phys. et rad. **8**, 23 (1947).
[11] M. H. Jellinek, E. Solomon, and I. Fankuchen, Ind. Eng. Chem. Anal. Ed. **18**, 172 (1946).
[12] S. H. Bauer, J. Chem. Phys. **13**, 450 (1945).

tering. However, the mathematical inversions involved in these treatments often appear unjustified by the quality of the available data, and the resultant distribution is not unique.

Although the general expressions of Guinier and of Vineyard[13] make no assumption regarding particle shape, it is frequently useful to assume a particle shape. Patterson[14] has derived a number of shape functions to describe the distribution of scattered intensity from particles of different shapes. In particular the spherical shape function is useful.

The spherical shape function was used by Guinier.[1] The relative intensity scattered by $M$ spherical particles is given by

$$I = MN^2\Phi^2(kR), \tag{1}$$

where $M$ = total number of scattering particles, $N$ = number of electrons per particle, $\Phi(kR)$ = spherical shape function, $= 3/(kR)^3[\sin(kR) - (kR)\cos(kR)]$, $k = 2\pi\epsilon/\lambda$, $\epsilon$ = scattering angle, $\lambda$ = wave-length of x-rays, $R$ = particle radius.

### INTERFERENCE EFFECT***

Equation (1) is derived from the amplitude of scattering from a single particle

$$E_1 = N\Phi(kR)\exp[2\pi ict/\lambda], \tag{2}$$

where $c$ = velocity of light, $t$ = time. Considering now the scattering from a second particle differing in phase from the scattering from the first particle by an amount $2\pi/\lambda[\bar{r}_m\cdot(\bar{s}-\bar{s}_0)]$, and summing over all particles, there results

$$E = \sum_m N\Phi(kR)\exp[2\pi i/\lambda\{ct - \bar{r}_m\cdot(\bar{s}-\bar{s}_0)\}], \tag{3}$$

where: $\bar{r}_m$ = vector from origin particle to $m$th particle, $\bar{s}$ = unit scattered vector, $\bar{s}_0$ = unit incident vector. The scattered intensity is then

$$I = E^*E = N^2\Phi^2(kR)\sum_m\sum_n \times\exp[2\pi i/\lambda\{\bar{r}\cdot(\bar{s}-\bar{s}_0)\}], \tag{4}$$

where $\bar{r} = \bar{r}_m - \bar{r}_n$. Following the procedure outlined by Gingrich,[15] Eq. (4) may be rewritten

$$I = MN^2\Phi^2(kR)\left[1 + \int_0^\infty 4\pi r^2\rho(r)\sin kr/krdr\right], \tag{5}$$

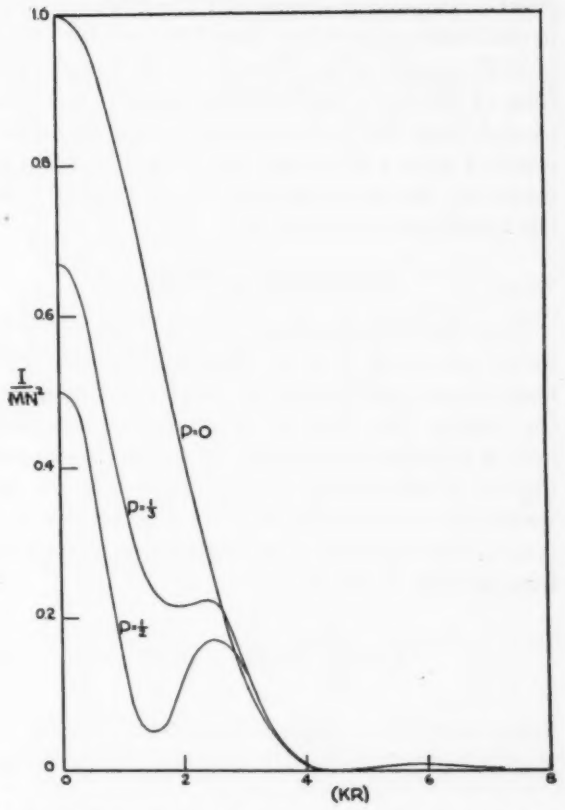



FIG. 2. Intensity *vs.* $(kR)$ from Eq. (6) for $P = 0, \frac{1}{3}, \frac{1}{2}$, showing interference maximum and first secondary shape function maximum.

where $\rho(r)$ is the density of particles at $r$. A discussion of the scattering from a monatomic fluid based on an equation of this form has been presented by Gingrich and Warren.[16] They assumed a density $\rho(r)$ of a form as shown in Fig. 1, and arrived at an equation of the form

$$I = MN^2\Phi^2(kR) \times[1 + P\{5\sin(2kR)/(2kR) - 6\Phi(2kR)\}], \tag{6}$$

where $0 < P < 1$ and is a parameter representing the closeness of packing. Equation (6) reverts to Eq. (1) for $P = 0$. Equation (6) is shown in Fig. 2 for $P = 0, \frac{1}{3}, \frac{1}{2}$. In addition to the interference maximum at $(kR) = 2.5$, the first of the series of secondary shape function maxima is seen at $(kR) = 5.8$.

Equation (1) is often approximated by the exponential:

$$I = MN^2\exp[-(kR)^2/5]. \tag{7}$$

[13] G. H. Vineyard, Phys. Rev. **74**, 1076 (1948).
[14] A. L. Patterson, Phys. Rev. **56**, 972 (1939).
*** Approach suggested by L. H. Lund.
[15] N. S. Gingrich, Rev. Mod. Phys. **15**, 90 (1943).
[16] N. S. Gingrich and B. E. Warren, Phys. Rev. **46**, 248 (1934).

A series expansion shows that the error involved in this assumption is less than 5 percent for values of $(kR)$ smaller than about 1.5. The logarithmic form of this equation has been used to determine particle size $(R)$ from the slope of the $\log I$ *vs.* $k^2$ graph. Figure 2 illustrates the radical error introduced by this procedure for values of $(kR)$ near the interference maximum.

## COLLIMATION ERROR

Since the maxima shown in Fig. 2 are not ordinarily observed, it is of interest to investigate those errors which may be effective in smearing the peaks. The two major causes of smearing appear to be non-uniformity of the specimens and imperfect collimation. Their effects may be described in terms of the intensity $I_p$ scattered to a point $P$ on the film. Considering first scattering from spheres of all sizes

$$I_p=\int_0^\infty [n(R)/M]IdR, \tag{8}$$

where $n(R)dR$ = number of particles having radius $R$ to $R+dR$, $I$ = intensity as given by theory for uniform particles, $M$ = total number of scattering particles.

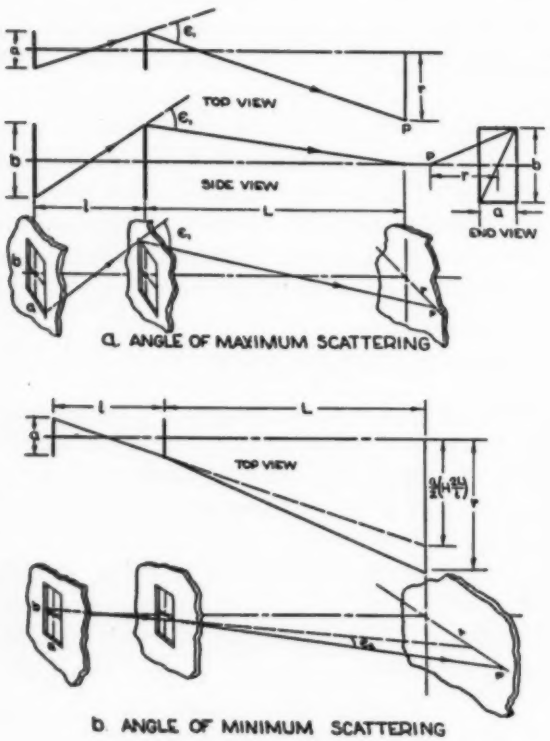



FIG. 3. Details of collimation system.

It should be noted that the limits of integration are infinite. Although knowledge of the specimen ordinarily allows establishment of a finite upper limit, the range may still be quite large, and the distribution $n(R)$ unknown. Treatments based on assumed distributions[8] may incur gross errors by neglecting a relatively few particles having extreme values of $R$. The effect may be minimized experimentally only by selecting the most uniform specimens available.

Considering now the error arising from imperfect collimation, an expression similar to Eq. (8) results

$$I_p=\int_{\epsilon_2}^{\epsilon_1} [n(\epsilon)/\zeta]Id\epsilon, \tag{9}$$

where $n(\epsilon)d\epsilon$ = power incident on particles of sample at proper directions to scatter to the point $P$ at angles $\epsilon$ to $\epsilon+d\epsilon$, $\epsilon_2$ = minimum scattering angle for ray reaching point $P$, $\epsilon_1$ = maximum scattering angle for ray reaching point $P$, $\zeta$ = total power incident on particles of sample.

The limits of integration are finite and readily determined from the slit geometry. The function $n(\epsilon)$ is not easily evaluated. However from physical considerations the assumption of a constant value for $n(\epsilon)$ through the range $\Delta\epsilon$ probably give rise to a slightly greater correction than most real distributions $n(\epsilon)$, where $\Delta\epsilon=\epsilon_1-\epsilon_2$. Then $n(\epsilon)\Delta\epsilon=\zeta$, and

$$I_p=\int_{\epsilon_2}^{\epsilon_1} I/\Delta\epsilon d\epsilon. \tag{10}$$

The error arising from the microphotometer slit employed has been evaluated separately and shown to be less than ten percent of the collimation error derived below. Other workers[8,17] have considered only slit height, neglecting the effect of slit width. Equation (10) may be evaluated using values of $I$ from Eqs. (1), (6), or (7). The use of Eq. (1) or (6) calls for numerical or graphical integration.

The limits $\epsilon_1$ and $\epsilon_2$ must also be evaluated. Figure 3 shows a collimation system of two identical slits of width $(a)$ and height $(b)$, sepa-

[17] M. H. Jellinek and I. Fankuchen, Ind. Eng. Chem. **37**, 198 (1945).

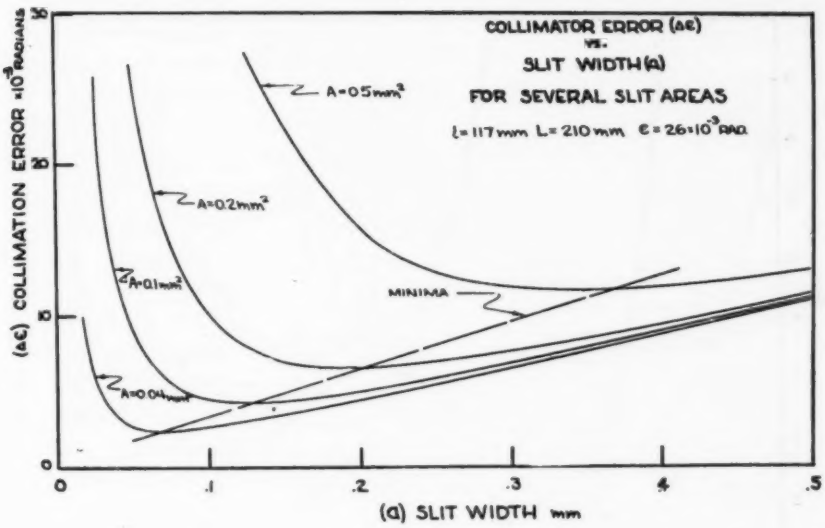



FIG. 4. Collimation error *vs.* slit width, showing minima.

rated by a distance ($l$), with the film a distance ($L$) from the second slit.

By assigning direction cosines to the incident and scattered rays in Fig. 3a, the angle of maximum scattering ($\epsilon_1$) may be expressed as a function of the geometry ($a$, $b$, $l$, $L$) and $r$, where $r$ is the distance from the center of the pattern to the point $P$. For $a$, $b$ and $r$ much smaller than $l$ and $L$, $\epsilon_1$ may be expressed,

$$\epsilon_1=[\xi^2a^2+2\xi a\epsilon+\epsilon^2+\xi^2A^2/a^2]^{\frac{1}{2}}, \qquad (11)$$

where $\xi=(2L+l)/2Ll$, $A=ab$, slit area, $\epsilon=r/L$. The angle of minimum scattering is seen at once from Fig. 3b to be

$$\epsilon_2=\epsilon-\xi a. \qquad (12)$$

A plot of $\Delta\epsilon$ as a function of ($a$) for selected values of ($l$, $L$, $A$, $\epsilon$) is shown in Fig. 4. Values of $l$ and $L$ selected are 117 mm and 210 mm; the value of $\epsilon$ is $26\times10^{-3}$ radian, to agree with experimental conditions. These curves exhibit minima which are significant as an indication of the slit dimensions yielding the minimum collimation error. The range of integration ($\Delta\epsilon$) is used as a measure of the collimation error for given $R$ and $\lambda$. More generally, $\Delta(kR)=[2\pi R/\lambda]\Delta\epsilon$ is a measure of the collimation error.

To investigate the effect of the collimation error in the region of the first secondary shape function maximum (at $kR=5.8$ in Fig. 2), the use of Eq. (1) for $I$ introduces little difference from the use of Eq. (6),

$$I_p=MN^2\int_{\epsilon_2}^{\epsilon_1}\Phi^2(kR)/\Delta\epsilon d\epsilon. \qquad (13)$$

This expression has been numerically integrated for several sets of experimental conditions. $I/MN^2$ from Eq. (1) is shown as the dotted curve $A$ in Fig. 5. The curves $B$, $C$, and $D$ represent $I_p/MN^2$ as given by Eq. (13) for $l=117$ mm, $L=210$ mm, $R=230$ angstroms and

$$\text{Curve } B\begin{cases}a=0.25 \text{ mm}\\ b=0.80 \text{ mm}\\ \lambda=8.32 \text{ angstroms}\end{cases}$$

$$\text{Curve } C\begin{cases}a=0.05 \text{ mm}\\ b=4.0 \text{ mm}\\ \lambda=8.32 \text{ angstroms}\end{cases}$$

$$\text{Curve } D\begin{cases}a=0.25 \text{ mm}\\ b=0.80 \text{ mm}\\ \lambda=1.54 \text{ angstroms.}\end{cases}$$

Curves $B$ and $D$ are drawn for near-optimum slit geometry as seen from the curve in Fig. 4 for a slit area of 0.2 mm². Curve $C$ is drawn for a long slit of the same area in which the collimation error is greater by a factor of three. Curve $B$ compared with $A$ shows a moderate smearing and a slight shift of the peaks towards smaller ($kR$). Curve $C$ shows more pronounced smearing and shifting as may be expected for the increased $\Delta\epsilon$ as seen in Fig. 4. Curve $D$ for the same slit condi-

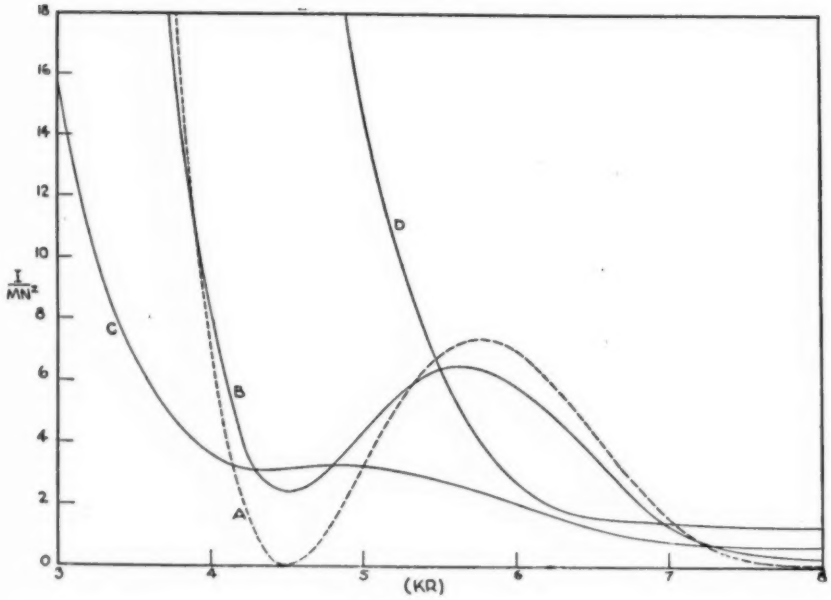



FIG. 5. Effect of collimation error as computed from Eq. (13), showing effect of optimum *vs.* long slits and soft *vs.* hard radiation.

tions as *A* but harder radiation shows a complete washing out of the peak.

## EXPERIMENTAL

Small angle x-ray scattering data and electron micrographs have been obtained on a number of samples of colloidal gold of radii fifty to four hundred angstroms. The sodium citrate reduction[18] was found by electron microscopy to yield quite uniform sols. Electron micrographs were obtained with an RCA model *B* instrument. The specimens were prepared for micrographing in the usual manner[19] by drying onto a collodion membrane. Calibration of the microscope was made by comparison with a shadow cast[20] Formvar replica[21] of a ruled optical grating.

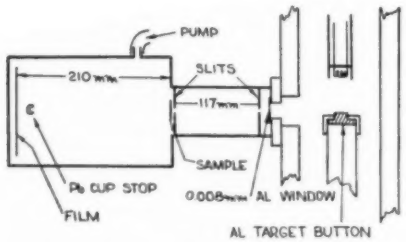



FIG. 6. Small angle scattering tube and camera.

The small angle scattering geometry is shown in Fig. 6. Eisenstein's demountable x-ray tube[22] was used as a source of soft radiation. An adjustable target was made using a replaceable button of soft aluminum. The $K\alpha$ radiation was excited by 4.1 kv electrons at a current of about 25 ma. The window and sample mount, both of 0.008-mm aluminum foil, together absorb 75 percent of the x-rays. The camera was evacuated by a mechanical pump. A Machlett A-2 tube was used as a source of copper radiation with a geometry identical with the aluminum camera geometry. In this case the film was covered with paper, and it was unnecessary to evacuate the camera. A Geiger detector was found useful for alignment of the slits and cup stop.

Small angle scattering data was recorded on Eastman Non-Screen Film, and the film densities were measured using a Leeds-Northrup recording microphotometer. This method of recording x-ray intensity was found to be linear for both copper and aluminum radiation over the entire range of recordable densities.

The use of filtering for monochromatization

[18] E. A. Hauser and J. E. Lynn, *Experiments in Colloid Chemistry* (McGraw-Hill Book Company, Inc., N. Y., 1940), p. 18.

[19] Zworykin, Morton, Ramberg, Hillier, and Vance, *Electron Optics and the Electron Microscope* (John Wiley and Sons, Inc., New York, 1945).

[20] R. C. Williams and R. W. C. Wycoff, J. App. Phys. **17**, 23 (1946).

[21] C. J. Burton, R. B. Barnes and T. G. Rochow, Ind. Eng. Chem. **34**, 1429 (1942).

[22] A. S. Eisenstein, Rev. Sci. Inst. **13**, 208 (1942).

has been justified by the validity of small angle data taken by other workers[23] using filtered radiation. In particular, an experiment by Lea[24] shows that aluminum filtered 4-kv aluminum radiation is about 99-percent equivalent K-series.

Three samples were selected for discussion. An electron micrograph of sample 1 is shown in Fig. 7. A particle size distribution obtained from direct measurement on this micrograph yields a radius of 232 angstroms plus or minus eleven percent. Figure 8 is a plot of the logarithm of the scattered x-ray intensity *vs.* $k^2$ for both copper $K\alpha$ and aluminum $K\alpha$ radiation as indicated. The copper curve exhibits no inflections, extending inward only to a value of $k^2=0.5\times10^{-4}$ radians$^2$/angstrom,$^2$ corresponding to $(kR)=1.6$. It should be remarked that inflections in the intensity curves located at the positions of the interference maximum and the first shape function maximum have been observed using copper radiation with 50 angstrom particles. This is expected as the collimation error $[2\pi R/\lambda]\Delta E$ is proportional to $R$. From the upward concavity of the copper radiation curve, it is expected that the use of the exponential approximation, Eq. (7), results

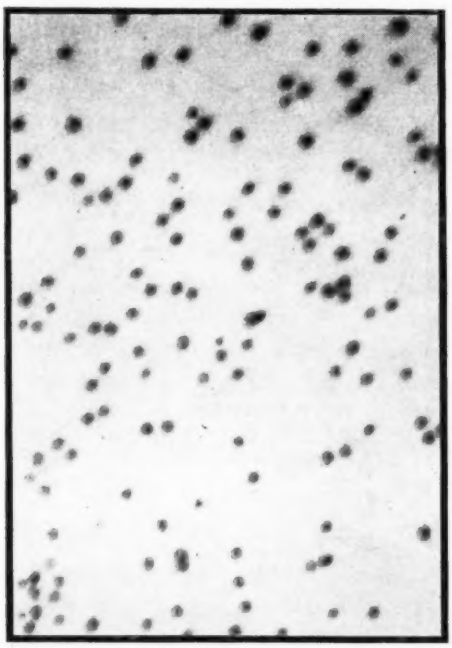


FIG. 7. Electron micrograph of sample 1 (232 Angstroms).

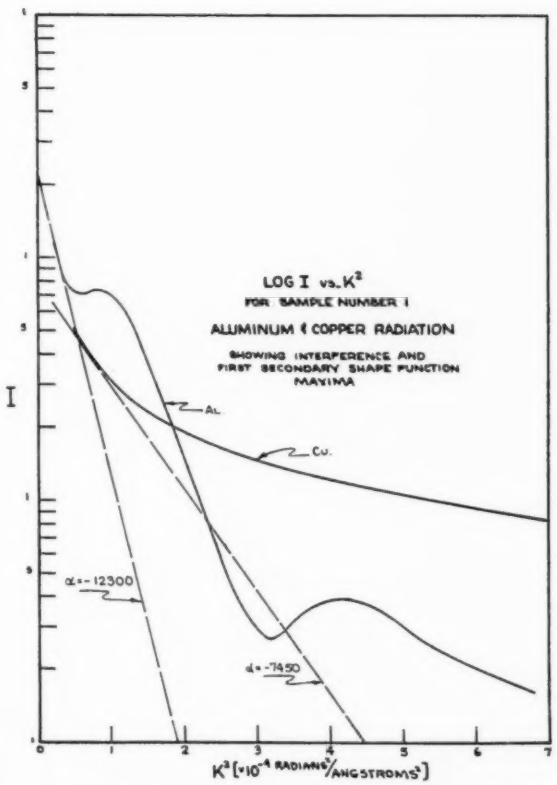



FIG. 8. Log $I$ *vs.* $k^2$ for sample 1, with copper and aluminum radiation. The interference maximum and first secondary shape function maximum appear on the aluminum curve.

in a somewhat low value for $R$ from the slope at $(kR)=1.6$. This is seen to be the case:

$$R=(-5\alpha)^{\frac{1}{2}},$$

where $\alpha=$ slope of log $I$ *vs.* $k^2$,

$$R=193 \text{ angstroms}.$$

The aluminum curve is readily extended to smaller $(kR)$ with the same geometry. From the aluminum curve in Fig. 8 the slope at $(kR)=1$ yields a value of

$$R=248 \text{ angstroms}.$$

For the region of $(kR)$ smaller than that corresponding to the interference peak, the slope is expected to be greater than that predicted by Eq. (1). Thus, a slightly large value of slope $\alpha$ and hence of $R$ may be expected.

The two peaks in Fig. 8 occur at $k^2=0.85$ and $4.2\times10^{-4}$ radians$^2$/angstrom.$^2$ A comparison with Fig. 2 identifies these peaks with the inter-

[23] H. H. Hubbell, Jr. thesis, Princeton University (1947).
[24] D. E. Lea, Amer. J. Roentgenology **45**, 614 (1941).

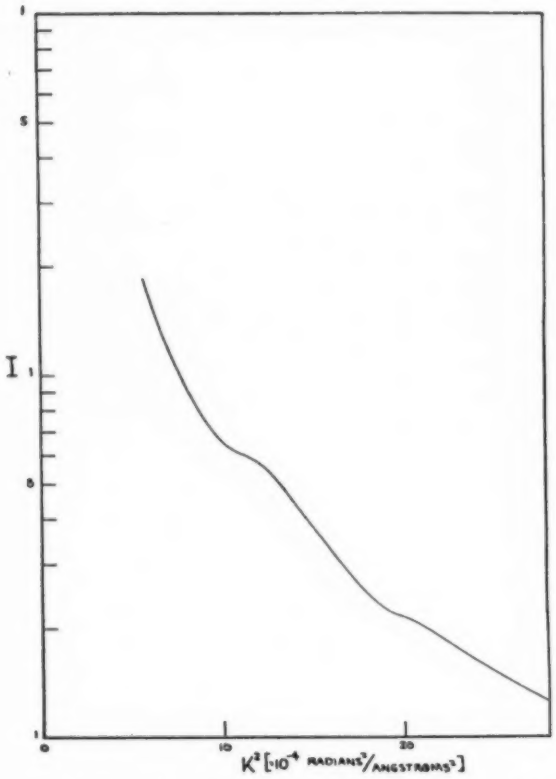



FIG. 9. Log $I$ *vs.* $k^2$ for sample 1 with aluminum radiation, showing second and third secondary shape function maxima.

ference maximum and shape function maximum at $kR=2.5$ and 5.8 respectively. From the positions of the first peak, a value of $R$ may be calculated

$$R=2.5/k=271 \text{ angstroms.}$$

From the second peak,

$$R=5.8/k=283 \text{ angstroms.}$$

Consideration of the peak shift to smaller ($kR$), shown in Fig. 5, as a result of collimation error and a similar shift resulting from non-uniformity in particle size leads one to expect high values of $R$ calculated from the uncorrected peak positions as above.

The intensity at the first peak in the experimental curve is $\frac{1}{3}$ of the zero angle intensity obtained by dash line extrapolation of the curve. Referring to Fig. 2, this is seen to agree well with the theoretically predicted intensity ratio for the parameters selected in writing Eq. (6). The relative intensity of the second experimental peak is somewhat higher than predicted. This is explained by the probable presence of a few small particles in the specimen which may scatter considerably at angles near the position of the shape function maxima.

Much longer exposure with larger optimum shaped slits reveals the second and third shape function maxima, as shown in Fig. 9. These maxima are observed only as inflections in the intensity curve due to the increased error from the larger slits required to get sufficient intensity to record them. The experimental peaks occur at about $k^2=11.5$ and $20.5\times10^{-4}$ radians²/angstrom.² Equation (1) predicts their position at $(kR)=9.1$ and 12.3, respectively. From the first peak in Fig. 9

$$R=9.1/k=260 \text{ angstroms.}$$

From the second peak:

$$R=12.3/k=272 \text{ angstroms.}$$

Again the inward peak shift has lead to high values for $R$.

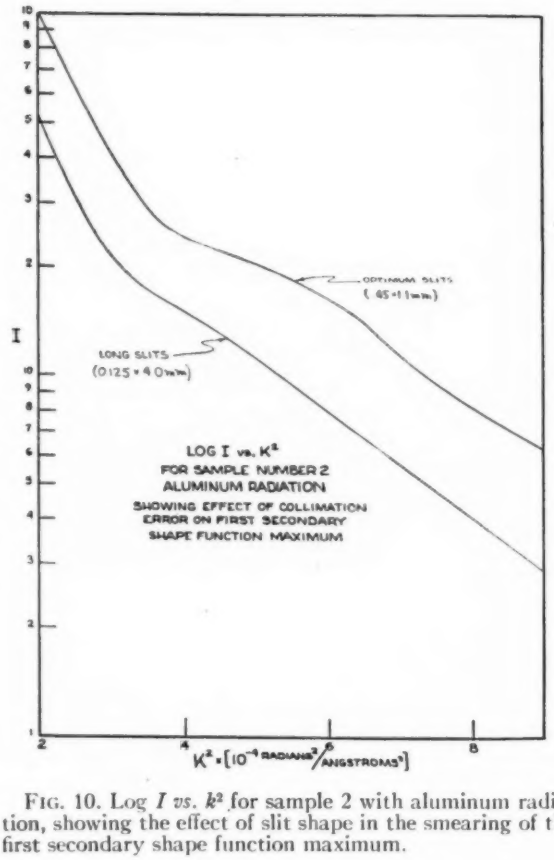



FIG. 10. Log $I$ *vs.* $k^2$ for sample 2 with aluminum radiation, showing the effect of slit shape in the smearing of the first secondary shape function maximum.

Sample 2 is almost identical with sample 1, having a mean particle radius of 232 angstroms plus or minus thirteen percent. An experimental verification of the effect of slit shape was made with this sample. Patterns were taken with optimum shape slits and with long slits of equal area. The resultant intensity in the region of the first secondary shape function maximum is shown in Fig. 10. The (0.45×1.1 mm) slits are seen from Fig. 4 to have a collimation error of less than half that for the long (0.125×4.0 mm) slits. The curve obtained using the optimum slits reveals a hump at about $k^2=6\times10^{-4}$ radians$^2$/angstrom.$^2$ This leads to

$$R=5.8/k=237 \text{ angstroms.}$$

The curve obtained using the long slits shows no measureable hump.

Sample 3 serves to demonstrate the extension of small angle scattering methods of particle size determination to larger particles. An electron micrograph of this sample is shown in Fig. 11. The particle radius from this micrograph is 380 angstroms plus or minus 16 percent. The increased non-uniformity of size and the increased particle radius [larger $\Delta(kR)$] both lead to the increased smearing of the peaks observed in Fig. 12. The slope analysis may be applied to the aluminum curve in Fig. 12 which extends in to $k^2=0.04$ radian$^2$/angstrom$^2$ or $(kR)=0.76$. From the slope of the tangent,

$$R=(-5\alpha)^{\frac{1}{2}}=387 \text{ angstroms.}$$

The copper curve on the other hand terminates at $k^2=0.9$ radians$^2$/angstrom,$^2$ corresponding to $(kR)=3.6$. This is well beyond the range of validity of Eq. (7). As the slope of the copper curve is increasing with decreasing $k^2$, the slope at the innermost end of the data may be expected to yield a low value for $R$;

$$R=(-5\alpha)^{\frac{1}{2}}=190 \text{ angstroms.}$$

It is readily shown from Fig. 3 that at angles smaller than $\xi a$, the unscattered beam is present. For the geometry used $\xi a$ equals about $2\times10^{-3}$ radians. For 380 angstrom particles and copper radiation this corresponds to $(kR)=3$ and for aluminum radiation to $(kR)=0.6$. Considering this limitation [$\epsilon>2\times10^{-3}$ radians] together with the condition prescribed, earlier for the application of Eq. (7) [$(kR)<1.5$] a general criterion may be formulated that for valid application of Eq. (7): $R<120\lambda$.

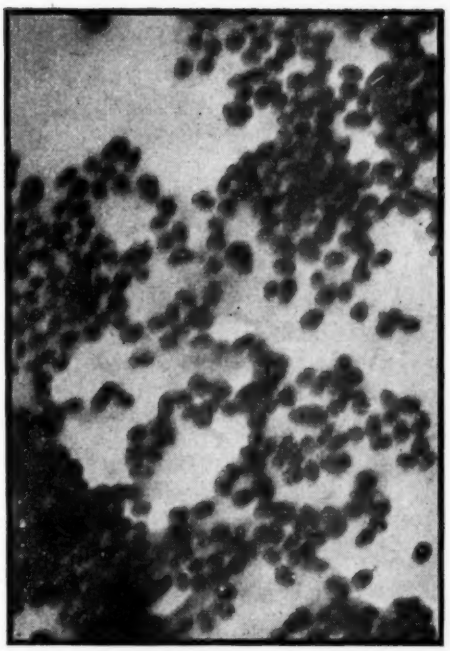


FIG. 11. Electron micrograph of sample 3 (380 Angstroms).

## CONCLUSIONS

A treatment based on earlier fluid scattering theories is developed for the interference between particles. The position and magnitude of the small angle scattered intensity maximum predicted by this treatment are confirmed by experiments using collodial gold specimens.

Several errors arising in small angle scattering measurements are discussed. The collimation error is evaluated taking account of slit width as well as slit height. An optimum height to width ratio is determined and experimentally verified.

The relative collimation error $[2\pi R/\lambda]\Delta\epsilon$ is reduced by using long wave-lengths (8.32 angstroms), as shown by the presence of pronounced intensity maxima. The positions and intensities of the first three secondary shape function maxima are checked with theory.

The usual method of small angle analysis is shown to be valid only for particles of radius less than 120 wave-lengths. The use of longer wave-lengths permits measurement on larger particles, extends the actual data spread for particles of a

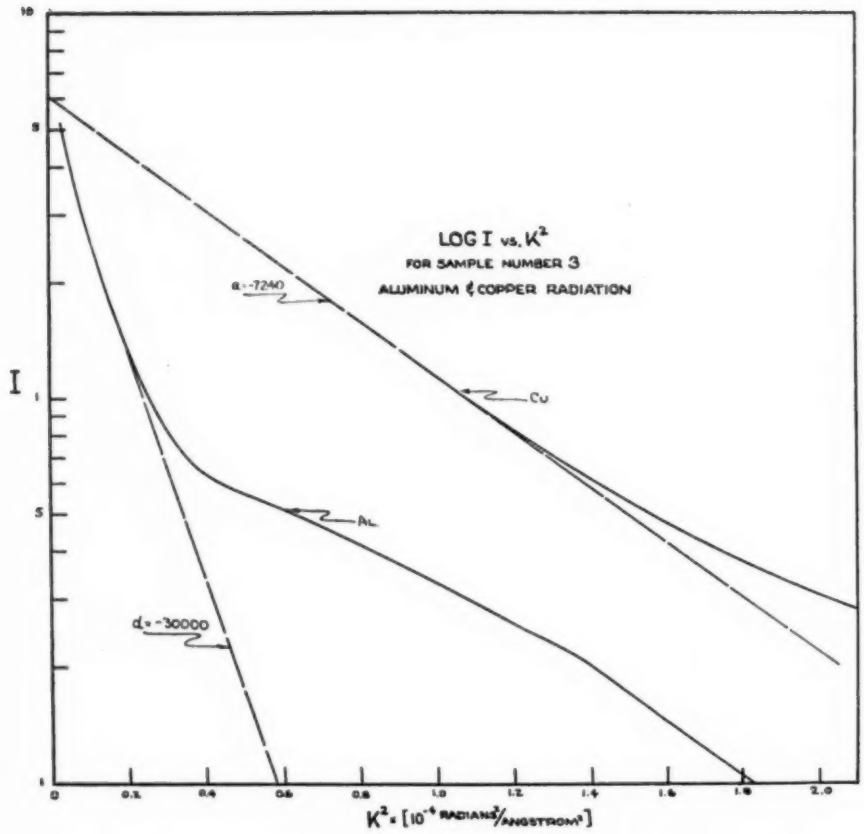



FIG. 12. Log $I$ *vs.* $k^2$ for sample 3 with both copper and aluminum radiation.

given size, and achieves improved electron microscope correlation.


## ACKNOWLEDGMENT

The writer expresses his appreciation to Professor N. S. Gingrich for suggestions leading to this study, and to Professor A. S. Eisenstein for guidance in this investigation. Grateful acknowledgment is made to the O. M. Stewart Fellowship Committee at the University of Missouri for the support which made this work possible.

# Aberration Correction with Electron Mirrors*

E. G. RAMBERG
*Radio Corporation of America Laboratories, Princeton, New Jersey*
(Received August 11, 1948)

Both the spherical and the chromatic aberration of electron microscope objectives may, in principle, be corrected with the aid of a uniform retarding field acting as a mirror. Such an arrangement has the drawback of requiring a conducting film in the ray path and the insertion of the specimen in a region of high field strength. The employment of concave electron mirrors with concentrated field distribution, forming a real image of approximately unity magnification, is free from this drawback. The formulas for spherical and chromatic aberration, presented in a form suitable for calculation, are applied to a characteristic electron mirror field of this type ($\Phi = C + \tanh(\sinh z)$). It is found that the aberration coefficients of the mirror are so large, however, that this method of aberration correction encounters serious practical difficulties.

IT is generally recognized that the spherical aberration of present-day electron-microscope objectives constitutes a limiting factor in the performance of electron microscopes. Proposals for correcting this defect have taken a number of different directions. Scherzer, in a recent critical examination,[1] lists the following procedures, partly devised by him:

1. The use of high frequency fields to vary the relative refractive power of the marginal and axial portions of the optical system during the transit of the electrons.
2. The introduction of balanced asymmetries into the lens system.
3. The insertion of controlled space charge.[2]
4. The insertion of axial electrodes.[3]
5. The placement of metal films into the ray path.
6. The utilization of electron mirrors.

Among all these methods, the employment of electron mirrors possesses advantages from the point of view of simplicity and the absence of essential asymmetries and obstructing materials from the critical imaging region. This justifies a closer examination of this possibility.

The fact that correction of spherical and chromatic aberration is possible with the aid of electron mirrors follows immediately from the characteristics of the uniform retarding field (Fig. 1).[4] If the field strength is $-\Phi' = -\Phi_i/d$ and the initial kinetic energy of the electrons is $e\Phi_i$, the equations of motion

$$\ddot{r} = 0; \quad \ddot{z} = e\Phi'/m = e\Phi_i/(md) \tag{1}$$

can be solved exactly, leading to the solution, for a ray leaving the point $z=0$, $r=0$ with an initial radial velocity component $(2e\Phi_r/m)^{\frac{1}{2}}$,

$$r = 2d\{[(\Phi_r/\Phi_i)(1-\Phi_r/\Phi_i)]^{\frac{1}{2}} \pm [(\Phi_r/\Phi_i)(1-\Phi_r/\Phi_i - z/d)]^{\frac{1}{2}}\}. \tag{2}$$

The apparent point of divergence of the reflected ray, in the plane $z=0$, is, hence, at a distance

$$-r/r' = 4d(1-\Phi_r/\Phi_i) = 4d(1-\theta^2) \tag{3}$$

from the plane; the prime indicates differentiation with respect to $z$ and $\theta$ is the (small) angle of inclination of the electron path at the point of origin (object point) relative to the field lines. It follows that, if this system is combined with an electron microscope objective with a focal length somewhat greater than $4d$ and a spherical aberration of $-4d\theta^3$—which is readily attainable with magnetic electron microscope objectives—the object being placed on a thin plane metallic film $A$ separating the region of uniform field in the mirror from field-free space in the lens, a system corrected for third-order spherical aberration is obtained (Fig. 1). Illumination would be provided from a source at the left, through the objective lens. Furthermore, if the same calculation is

* Presented at the Second Annual Symposium on Applied Mathematics, Cambridge, Mass., July 29–31, 1948.

[1] O. Scherzer, Optik. 2, 114–132 (1947).

[2] D. Gabor, *The Electron Microscope* (Hulton Press, London, 1944).

[3] D. Gabor, Nature 158, 198 (1946); J. W. Dungey and C. R. Hall, Proc. Phys. Soc. London 59, 828–843 (1947).

[4] Zworykin, Morton, Ramberg, Hillier, and Vance, *Electron Optics and the Electron Microscope* (John Wiley & Sons, Inc., New York, 1945), p. 630 (cited below as EOEM).

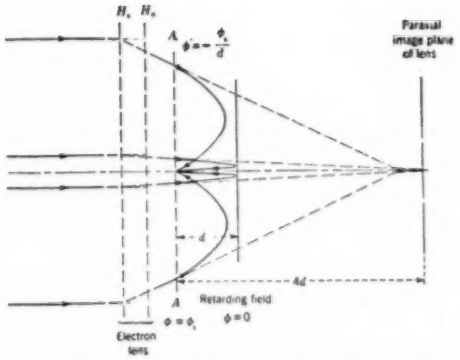



FIG. 1. Correction of aberrations of an electron lens by combination with a retarding field acting as an electron mirror.

carried out for an initial kinetic energy $e(\Phi_i+\Delta\Phi)$, it is found that the point of divergence is shifted away from the plane $z=0$ by $4d\Delta\Phi/\Phi_i$. Since the chromatic shift of focus of a thick magnetic lens is somewhat smaller than $f\Delta\Phi/\Phi_i$, where $f$ is the focal length of the lens, the system described promises to correct spherical and chromatic aberration simultaneously.

Unfortunately, from a practical point of view the above system is less attractive. The necessary conducting film intersects the ray paths in an optically active region and is subject to very high fields. The specimen, similarly, is in a high field region, making its insertion and manipulation extremely difficult. In brief, it seems that the system described has few, if any, advantages and many serious disadvantages as compared with objectives corrected with conducting films.

To make an electron mirror correcting system practical it seems essential that the object be placed in field-free space, making the insertion of a conducting film unnecessary and facilitating manipulation of the specimen. Furthermore, it appears desirable to let the mirror produce a real image fairly close to the object, so that disturbances which may be introduced by the latter have a relatively small effect on image resolution. Figure 2 shows, schematically, an arrangement which fulfills these conditions. The illuminating beam is separated, in the example shown, from the imaging beam by a transverse magnetic field; the distortion introduced by the transverse field may eventually be corrected with the aid of symmetrizing screws in the projector.[5] The object is placed a small distance from the focal plane of the objective and the "concave" mirror[6] voltage adjusted to form an image of it in the focal plane. It is seen that the illuminated area of the object is determined by the separation of the object and focal plane and by the relative aperture of the objective. The reflected rays pass again through the object, which presents least obstruction if it is freely suspended in the object support openings.

A mirror field which satisfies the requirements of smooth variation and relatively rapid cut-off is represented by the axial potential

$$\Phi=\tfrac{1}{2}(\Phi_1+\Phi_0)+\tfrac{1}{2}(\Phi_1-\Phi_0)\tanh(\sinh z) \quad (4)$$

to which the off-axis potential is related by

$$\varphi(z,r)=\sum_{n=0}^{\infty}\frac{(-1)^n}{(n!)^2}\Phi^{(2n)}\left(\frac{r}{2}\right)^{2n}. \quad (5)$$

Close to $z=0$ this field resembles that between two coaxial equidiameter cylinders with a diameter 2.63 measured in the same units as $z$.[7] For values of $z$ in excess of 2, on the other hand, the potential rapidly reaches a constant value and the field-determining electrodes approach the axis closely, so that in this region changes in the electrode configurations have little effect on the action of the field. This is evident from the equipotential pattern shown in Fig. 3, which was determined with the aid of Eq. (5).

The fact that the ray paths exhibit an infinite slope at the reversal point even in the paraxial approximation makes it desirable to employ a parametric representation of the paths. With the time as parameter the path equations become the usual equations of motion:

$$\ddot{r}_s=\frac{e}{m}\frac{\partial\varphi_s}{\partial r}=-\frac{e}{2m}r_s\Phi_s''+\frac{e}{16m}r^3\Phi^{IV}, \quad (6)$$

$$\dot{z}_s=\pm\left(\frac{2e\varphi_s}{m}-\dot{r}^2\right)^{\frac{1}{2}}=\pm\left(\frac{2e\Phi_s}{m}-\frac{er^2\Phi''}{2m}-\dot{r}^2\right). \quad (7)$$

Here the subscript $s$ denotes the second approximation, corresponding to the retaining of the third-order aberration terms. If $r(t)$, $z(t)$ is the

[5] See J. Hillier and E. Ramberg, J. App. Phys. **18**, 48–71 (1947).

[6] It may be noted that the "convex" electron mirrors which have been studied were found to have the wrong sign of spherical aberration to bring about correction (see e.g., EOEM, p. 631).

[7] See S. Bertram, Proc. I. R. E. **28**, 418–420 (1940).

paraxial solution for an electron leaving the axial object point $z_0$ at the time $t_0$ with an initial radial velocity component $\dot{r}_0$:

$$\ddot{r}=-\frac{er}{2m}\Phi''; \quad t-t_0=\int_{z_0}^{z}(2e\Phi/m)^{-\frac{1}{2}}dz \quad t<t_r$$

$$t-t_r=-\int_{z_r}^{z}(2e\Phi/m)^{-\frac{1}{2}}dz \quad t>t_r. \quad (8)$$

Here $t_r$ is the paraxial reversal time, corresponding to the arrival of the electron at $z_r(\Phi(z_r)=0)$. It is assumed that the electron is incident on the mirror in the positive direction along the $z$-axis. Near the point of reversal, $\Phi(z)=\Phi'(z_r)(z-z_r)$, so that the evaluation of the integral does not cause any difficulty. The derivation of the second approximation from this solution is then given by

$$\Delta r=r_s-r, \quad \Delta z=z_s-z,$$

$$\Delta\ddot{r}+\frac{er}{2m}\Phi''\Delta r=R(t)=-\frac{e\Phi'''}{2m}r\Delta z+\frac{e\Phi^{IV}}{16m}r^3; \quad (9)$$

$$\Delta\dot{z}=\pm\left(\frac{2e\Phi}{m}\right)^{\frac{1}{2}}\left(\frac{\Phi'}{2\Phi}\Delta z-Q(t)\right), \quad (10)$$

$$Q(t)=\frac{\Phi''}{8\Phi}r^2+\frac{m\dot{r}^2}{4e\Phi}.$$

The solution of the equation for $\Delta r$ can be carried out by standard methods[8] provided that $\Delta z(t)$ has been found. Thus, for $t=t_i$, corresponding to the time at which the paraxial electron has returned to the axis,

$$\Delta r(t_i)=-(M/\dot{r}_0)\int_{t_0}^{t_i}R(\tau)r(\tau)d\tau, \quad (11)$$

where $M$ is the magnification.

The determination of $\Delta z$ is somewhat more complex. For $t<t_r$ Eq. (10) is readily integrated to yield

$$\Delta z=-(2e\Phi/m)^{\frac{1}{2}}\int_{t_0}^{t}Q(t)dt. \quad (12)$$

Since the integrand becomes infinite as $t\to t_r$, it is advisable to carry out the integration by parts suggested by following form of Eq. (12):

$$\Delta z=\left(\frac{2e\Phi}{m}\right)^{\frac{1}{2}}\left\{-\int_{t_0}^{t_1}Q(t)dt+\int_{t_1}^{t}\frac{2\Phi}{\Phi'}Q(t)\frac{d}{dt}\left(\frac{m}{2e\Phi}\right)^{\frac{1}{2}}dt. \quad (13)\right.$$

This leads to the expression, for $t\lesssim t_r$:

$$\Delta z=\frac{2\Phi}{\Phi'}Q(t)\mp\left(\frac{2e\Phi}{m}\right)^{\frac{1}{2}}\times\left\{\int_{t_1}^{t}\left[\left(\frac{\Phi'''}{4\Phi'}-\frac{\Phi''^2}{4\Phi'^2}\right)r^2-\frac{m\Phi''}{2e\Phi'^2}\dot{r}^2\right]dt-\left(\frac{m}{2e\Phi_1}\right)^{\frac{1}{2}}\left[\Delta z_1-\frac{2\Phi_1}{\Phi_1'}Q(t_1)\right]\right\}. \quad (14)$$

The expression with the lower sign, for $t>t_r$, is obtained with the aid of the requirement that both $\Delta z$ and $\Delta\dot{z}$ be continuous at $t=t_r$. Equation (14) contains only finite terms and is employed in a range of $t(t_1<t<t_2)$ about $t_r$ in which $\Phi'$ has relatively large values.[9] For $t>t_2$, finally,

$$\Delta z=\left(\frac{\Phi}{\Phi_2}\right)^{\frac{1}{2}}\Delta z_2+\left(\frac{2e\Phi}{m}\right)^{\frac{1}{2}}\int_{t_2}^{t}Q(t)dt. \quad (15)$$

In terms of $\Delta r(t_i)$ and $\Delta z(t_i)$, the deviation of the ray in the paraxial image plane becomes

$$\Delta r(z_i)=\Delta r(t_i)+\Delta z(t_i)\dot{r}(t_i)(2e\Phi_i/m)^{-\frac{1}{2}}. \quad (16)$$

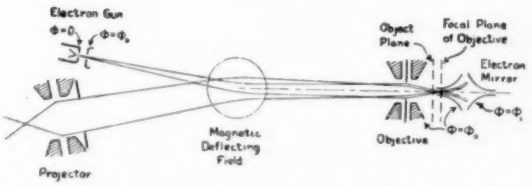



FIG. 2. Correction of aberrations of an electron microscope objective by combination with a "concave" electron mirror.

[8] See e.g., E. P. Adams, *Smithsonian Mathematical Formulae and Tables of Elliptic Functions* (Smithsonian Institution, Washington, 1922), Formula 8.410.

[9] It should be noted that the constant coefficient of $(2e\Phi/m)^{\frac{1}{2}}$ in Eq. (16.85) of EOEM is in error; this equation should, accordingly, be replaced by Eq. (14) of this paper, generalized to include the effect of a magnetic field. A similar change to

$$\Delta z=-\Delta\Phi\left\{\frac{1}{\Phi'}\pm\left(\frac{2e\Phi}{m}\right)^{\frac{1}{2}}\times\left[\int_{t_1}^{t}\frac{\Phi''}{\Phi'}dt-\left(\frac{\Phi}{\Phi_1}\right)^{\frac{1}{2}}\left(\frac{1}{\Phi_1'}+\frac{\Delta z_1}{\Delta\Phi}\right)\right]\right\}, \quad t\lesssim t_r$$

is required in Eq. (16.125) of EOEM.

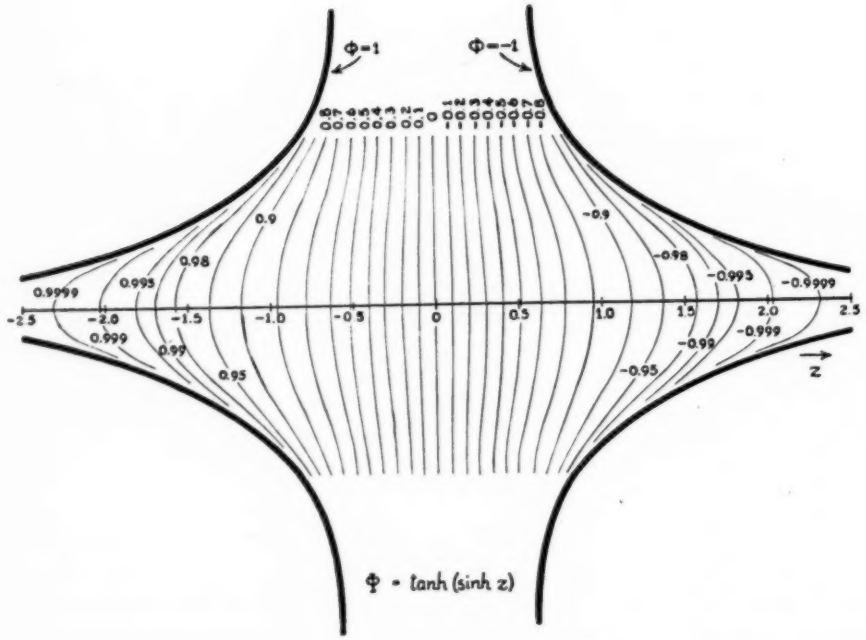



FIG. 3. Potential distribution in the mirror field $\Phi = \tanh(\sinh z)$.

To make the combined spherical aberration of the mirror and objective lens equal to zero it is necessary that the quantity $\Delta r(z_i)/r_i'^3$ be equal to the aberration constant $Cf$ of the objective. The actual evaluation yielded:

$\Delta r(z_i)/r_i'^3 = 7560$ units for $z_i = -8.55$, $M = -1$, $\Phi_1/\Phi_0 = -0.219$,
$\Delta r(z_i)/r_i'^3 = 602$ units for $z_i = -3.74$, $M = -1$, $\Phi_1/\Phi_0 = -0.169$,
$\Delta r(z_i)/r_i'^3 = 67$ units for $z_i = -2.15$, $M = -1$, $\Phi_1/\Phi_0 = -0.128$.

It is seen that the aberration decreases very rapidly as the mirror electrode is made less negative and that the aberration of the mirror is extremely large, so that, to balance the aberration of normal electron microscope objectives (which is of the order of a fraction of a centimeter) the mirror would have to be scaled down to unreasonably small dimensions. The last case listed represents the closest permissible approach of object and image point to the center of the mirror consistent with the requirement that the object be in practically field-free space, so that the aberration constant of 67 units may also be regarded as a minimum. Hence the scale unit in Fig. 3 would have to be reduced to a fraction of 0.1 mm (!) to lead to a mirror which would correct a normal high-quality electron microscope objective. It is interesting to note that the chromatic aberration of the same system[10] is given by

$$\Delta z = 70.8\Delta\Phi/\Phi_0$$

so that the balancing of the chromatic aberration of the objective (given by about $0.8f\Delta\Phi/\Phi_0$) would demand a scale reduction factor of the same order.

The very large values of the aberration constants obtained in the present example may be regarded as characteristic of concave electron mirrors; values of similar magnitude have been obtained for electron mirrors with a more gradually decreasing field. The neglect of the relativistic effect is certainly of secondary importance in electron microscope applications, particularly since it is least in the critical portion of the mirror. It is thus seen that the correction of electron microscope objectives by electron mirrors, employed in the manner described, meets serious practical difficulties.

[10] Calculated by the methods outlined in EOEM, pp. 575–577, with the correction mentioned in footnote 9.

# Sweep Frequency Ionosphere Equipment

PETER G. SULZER
*Department of Electrical Engineering, The Pennsylvania State College, State College, Pennsylvania*
(Received August 12, 1948)

A new sweep frequency ionosphere sounding equipment is described. The equipment records ionosphere virtual height as a function of frequency over the range from 1 to 25 megacycles. Special features of the device are high power output, good receiver sensitivity, and anti-jamming circuits.

## INTRODUCTION

IT is the purpose of this paper to describe the sweep frequency ionospheric recorder recently completed at The Pennsylvania State College. As normally used, the equipment sweeps the frequency range from 1 to 25 megacycles, recording ionosphere virtual height as a function of frequency. The pulse method of measurement is used, as first described by Breit and Tuve,[1] and as later used in radar.

Although such equipments have been described before,[2,3] the new recorder is believed to be of particular interest because of the following features:

(a) high power output, 10-kilowatt peak;
(b) excellent receiver sensitivity; signals of less than one microvolt are recorded;
(c) antijamming circuits, which provide good discrimination between pulsed and continuous-wave (cw) signals, thus preventing most of the cluttering of the records by communications services operating within the frequency range of the equipment.

## SYSTEM REQUIREMENTS

The most difficult requirement to meet in a transmitter-receiver system operating over such a wide frequency range is the tracking of the receiver with the transmitter. Other requirements which must be met at the same time are power output, receiver sensitivity, receiver band width, transmitted pulse shape, frequency stability, and height measurement accuracy.

Tracking was accomplished, as shown in Fig. 1, by controlling the operating frequencies of both the transmitter and the receiver by means of a common oscillator whose frequency is varied from 31 to 55 megacycles. In the transmitter, the output of this oscillator is heterodyned with that of a pulsed 30-megacycle oscillator, and the resulting frequency difference, which varies from 1 to 25 megacycles, is amplified and applied to the transmitting antenna. In the receiver, the output of the common oscillator is heterodyned with incoming signals, and the resulting constant

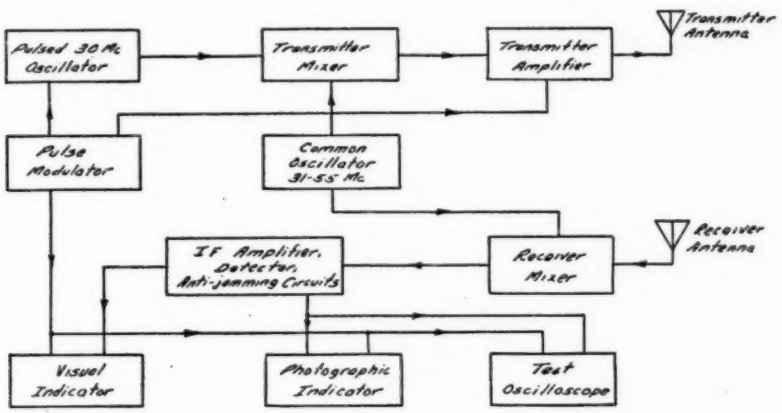



FIG. 1. System block diagram.

[1] G. Breit and M. A. Tuve, "A test for the existence of the conducting layer," Phys. Rev. 28, 554 (1926).
[2] T. R. Gilliland, "Field equipment for ionospheric measurements," J. Research. Nat. Bur. Stand. 26, 377 (1941).
[3] P. G. Sulzer, "Ionosphere measuring equipment," Electronics 19, No. 7, 137 (1946).

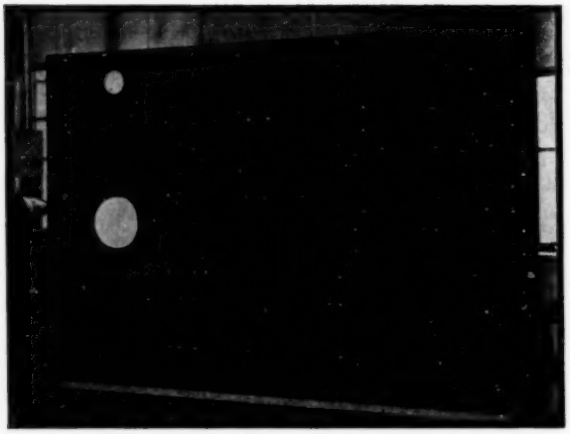


FIG. 2. Front-view photograph of the equipment.

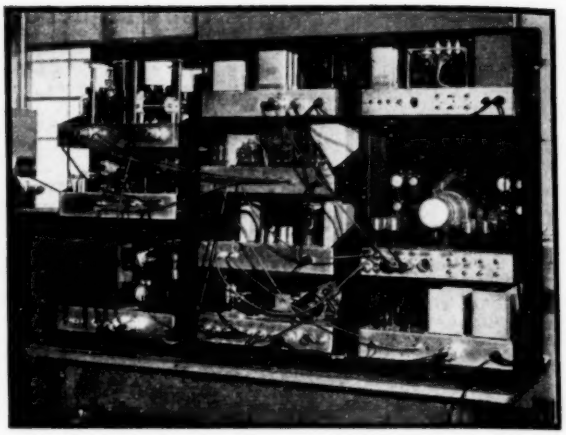


FIG. 3. Rear-view photograph of the equipment.

frequency difference of 30 megacycles is amplified and detected.

Thus the main requirement of tracking is easily met. It has been found that the equipment can be considerably simplified if the signal-frequency parts of the transmitter and receiver are not tuned. With modern high gain tubes it is possible to obtain sufficient power and sensitivity over the operating range of 24 megacycles with band-pass coupling circuits. By operating the common oscillator at a very low power level, good frequency stability can be obtained. The other requirements are discussed below.

The following description of the equipment is divided into three parts: the transmitter, receiver, and indicators. Front and rear-view photographs of the equipment appear in Figs. 2 and 3, respectively; Fig. 4 is a photograph of the recording camera mounted on its control and power box.

## TRANSMITTER

As explained above, the operating frequency of the transmitter is controlled by the common oscillator, $V_1$ of Fig. 5. The frequency of this oscillator is varied by an inductive tuning system which employs a tapered two-wire transmission line wrapped around a rotating drum. The transmission line rides against two spring contacts, varying its active length. The taper is set so that a constant rate of frequency change is obtained, which produces a linear frequency scale in the indicators. The oscillator is followed by two buffer stages, $V_2$ and $V_3$, which help to prevent feedback from the transmitter to the receiver.

The output of $V_3$ is carried to the transmitter intermediate amplifier, Fig. 6, by means of a coaxial cable. This unit contains $V_6$, the pulsed 30-megacycle oscillator; $V_7$, the transmitter mixer; and three stages of amplification, $V_8$, $V_9$, and $V_{10}$. The oscillator is controlled by a 150-volt positive pulse obtained from the modulator, which is described below. Care was taken to make the tuned circuit, which controls the frequency of the oscillator, identical with those

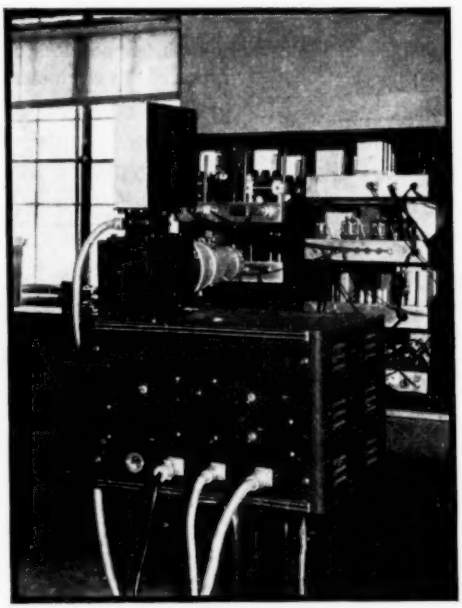


FIG. 4. Photograph of the recording camera and indicator.

used for coupling in the intermediate frequency amplifier, so that like temperature changes in both units would produce the same effect, and tracking would not be impaired.

The mixer, $V_7$, was adjusted for minimum distortion, since there is no way of discriminating between the fundamental frequency and harmonics in the remainder of the transmitter. This is followed by three stages of amplification, $V_8$, $V_9$, and $V_{10}$, which employ conventional video amplifier compensation to obtain the required band width of 24 megacycles.* The output of the last stage is 60 volts r.m.s. during the pulse interval. It should be noted that the last stage is pulsed, which results in a very substantial saving of power.

The remainder of the radiofrequency part of the transmitter, the power amplifier, appears in Fig. 7. It consists of three screen-pulsed stages of class-A video amplification using type 715-B tetrodes. The first stage, $V_{11}$, is a conventional amplifier with the load connected in the plate circuit. The voltage gain of this stage is 7, which is unusually high for such a great band width. This is a direct result of pulsed class-A operation, which permits high plate current, and consequently a very high value of grid-plate transconductance. The second stage, $V_{12}$, is a plate-cathode phase inverter, which provides balanced output for driving the final push-pull stage. Low load resistors were used in the phase inverter so that high frequency compensation would not be necessary. It was felt that the phase-response of such compensating circuits might interfere with the operation of the push-pull stage. The final stage, $V_{13}$ and $V_{14}$, connected in push-pull, provides a peak output of 10 kilowatts into a 600-ohm resistive load. The output is constant within 40 percent over the frequency range from 1 to 25 megacycles.

The pulse modulator, Fig. 8, generates the pulses necessary for keying the various stages of the transmitter and triggering the indicator circuits. It consists of a blocking oscillator, amplifiers, and cathode-follower coupling stages.

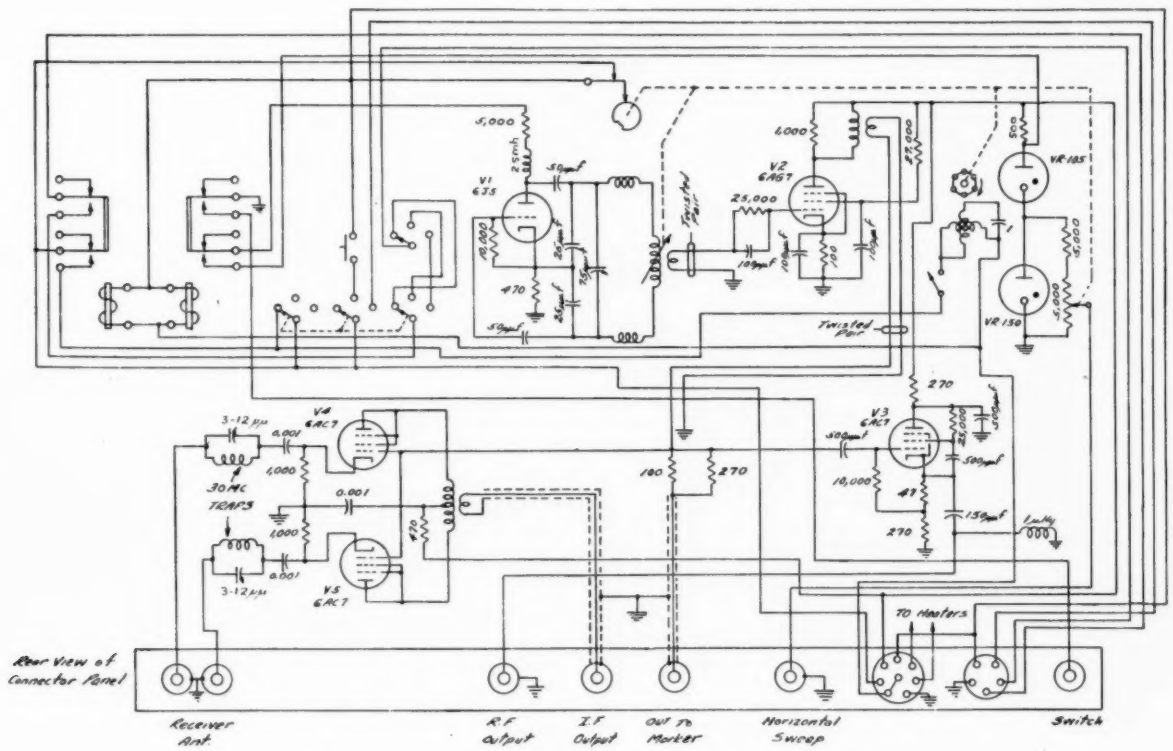



FIG. 5. Schematic diagram of the common oscillator unit.

* The transmitter amplifiers were designed and constructed by Gordon H. T. Lee.

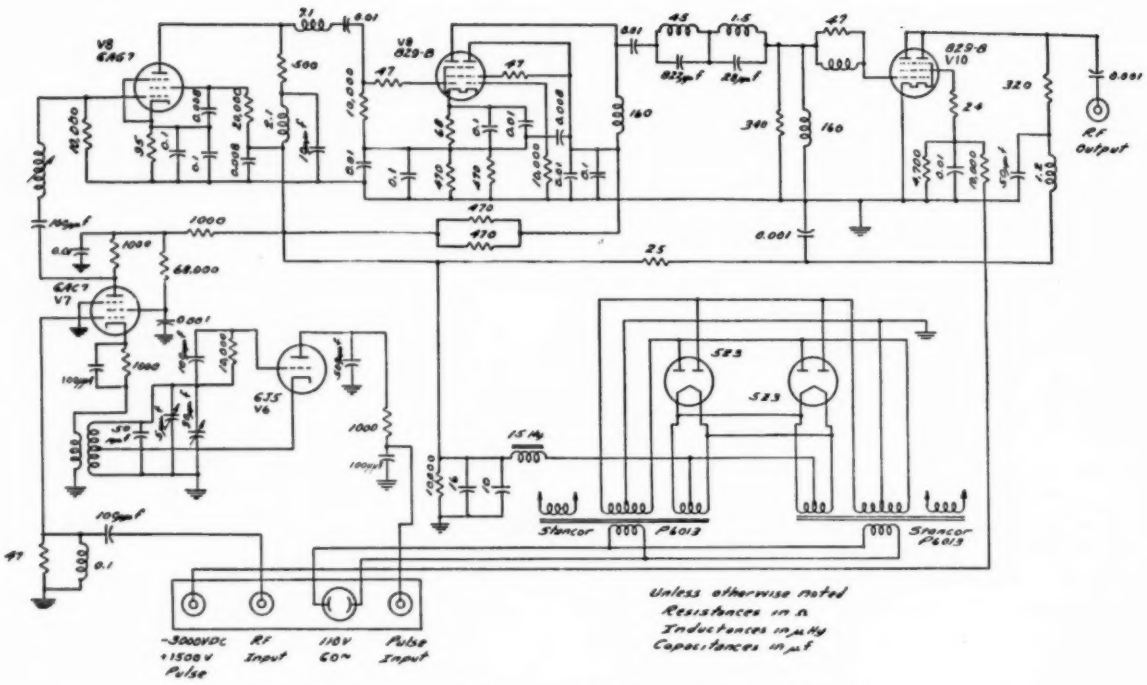



FIG. 6. Schematic diagram of the intermediate amplifier.

The blocking oscillator[4] and its synchronizing amplifier, $V_{15}$, are used to generate short negative pulses synchronized to the power line frequency, or to some external source. The normal pulse rate is 60 per second, with provision for operating at any frequency between 20 and 100 per second.

The output of the blocking oscillator is used to drive the indicator circuits, and is also used to actuate a trigger circuit,[5] $V_{16}$, which develops both positive and negative rectangular pulses of variable duration. The pulse duration is variable from 20 to 200 microseconds, but is normally set at 50 microseconds.

The negative pulse is used to drive an amplifier and cathode follower, $V_{17}$, which provides a positive 150-volt pulse. Since this voltage is used to key the transmitter oscillator, it is necessary that its amplitude remain constant during the active interval. Otherwise the resulting frequency modulation would spread some of the transmitter power outside the receiver pass-band.

The positive pulse from the trigger circuit is used to drive $V_{18}$, an amplifier which is normally biased to cut-off. The negative plate voltage pulse from this stage is inverted by means of a transformer, and applied to the grids of three cathode followers, $V_{19}$, $V_{20}$, and $V_{21}$. These tubes are provided with negative cathode returns so that the quiescent cathode voltage is $-200$, which is sufficient to cut off the stages being controlled. During the pulse the cathode voltage rises to $+1400$. Cathode followers were used as output stages to provide good pulse shape and to permit isolation between the various pulsed stages. The pulse top is flat within 5 percent, and the rise time from bottom to top is only 2 microseconds.

## RECEIVER

The receiver push-pull mixer, $V_4$ and $V_5$, appears in Fig. 5. Triode-connected tubes were chosen because of their low equivalent noise resistance, which permits low noise figure. Local-oscillator voltage is applied between the control grids, which are connected in parallel, and ground. This method of oscillator injection was used because only a small amount of oscillator power was available. The antenna connection is

[4] L. N. Ridenour (Ed.), *Radar System Engineering* (McGraw-Hill Book Company, Inc., New York, 1947), p. 502.
[5] See reference 4, p. 497.

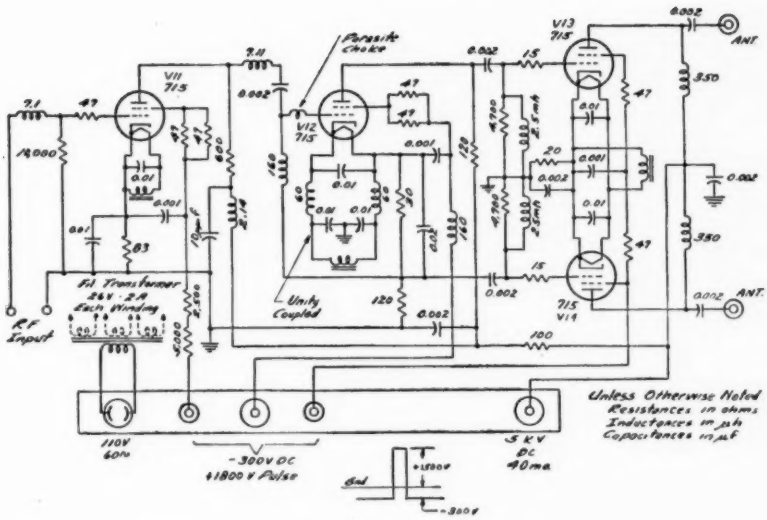



FIG. 7. Schematic diagram of the power amplifier.

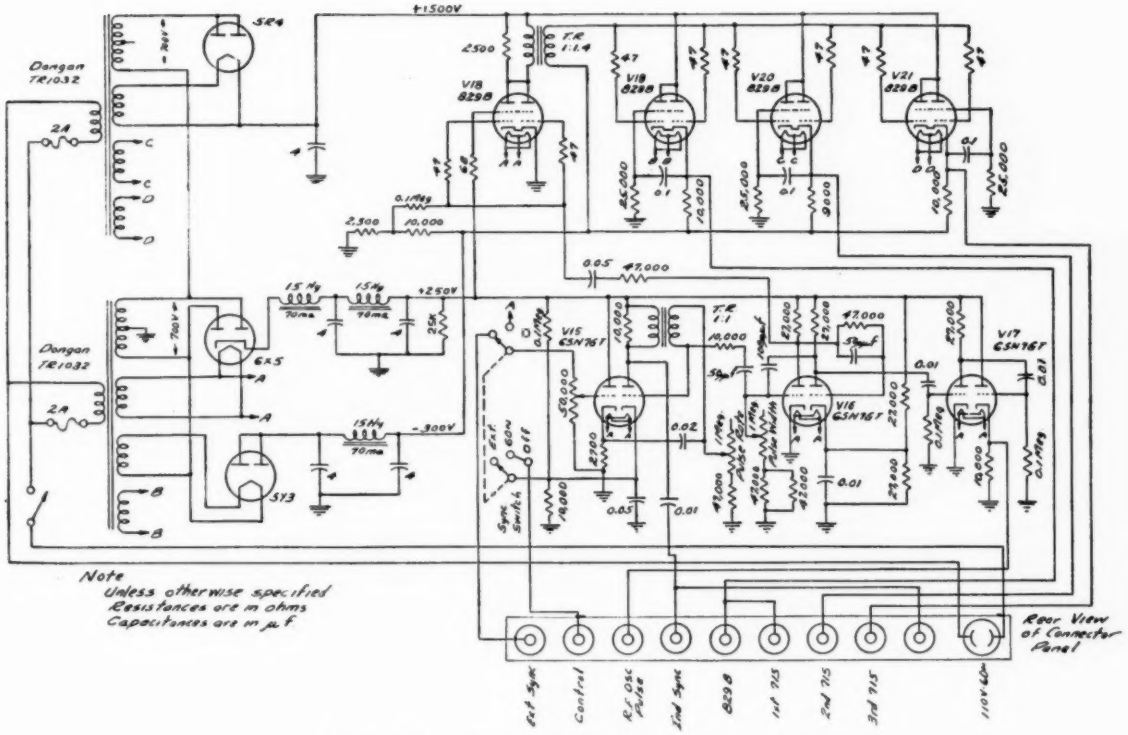



FIG. 8. Schematic diagram of the pulse modulator.

made to the cathodes of the tubes, which are connected in push-pull.

The remainder of the receiver appears in Fig. 9. This includes the intermediate frequency (IF) amplifier, detector, video amplifier, differentiator,[6] and instantaneous automatic gain control (IAGC)[7] circuits.

[6] See reference 4, p. 458.
[7] See reference 4, p. 459.

An intermediate frequency of 30 megacycles was dictated by the type of tracking system used. With this high a frequency, a band pass of 40 kilocycles, which is about the maximum that can be tolerated, is difficult to obtain unless double conversion is used. Double conversion is undesirable because of the many spurious signals obtained unless very good shielding is used. It

was found, however, that with careful design of the coils and associated circuits a $Q$ of 260 could be obtained. With six such tuned circuits the half-power band-pass is just 40 kilocycles. Thus, an amplifier with 5 stages, $V_{22}$ through $V_{26}$, was chosen. The IF amplifier has a maximum voltage gain of $10^6$, which is sufficient amplification after the mixer to register thermal noise from the antenna.

A IN34 crystal diode, $V_{29}$, was chosen for the second detector because it is very compact and does not have appreciable contact potential, which would interfere with the operation of the IAGC system.

It was found necessary to provide an IAGC system for the following three reasons:

(1) As the equipment sweeps through the frequency range 1–25 megacycles, certain strong signals overload the receiver. Since these signals are usually grouped in bands, the result is that large portions of the record are obliterated.

(2) Ionospheric absorption is low during the dark hours. The resulting strong signals produce the same effect as (1).

(3) Very strong reflected pulses will produce a wide trace on the record unless the receiver gain is decreased.

Since the gain control is to be operated during the pulse interval, it is necessary that very short time constants be used in the circuits. The system finally used employs direct coupling from the crystal mixer to the grids of three of the stages, which use remote cut-off pentodes.

As an additional aid to discrimination between pulsed and CW signals it was found helpful to incorporate a short time constant (differentiator) circuit between the first and second video stages, $V_{27}$ and $V_{28}$. It was necessary to follow this circuit with a clipper, $V_{30}$, to prevent overshoot. The remainder of the video system is conventional; it consists of a pentode stage followed by a triode phase inverter and cathode follower.

## INDICATORS

Two indicators are provided, one for photographic recording and one for visual observation. The photographic indicator uses a type 5FP7 cathode-ray tube. The pattern obtained is rectangular, with virtual height as the ordinate and frequency as the abscissa. Photographic records

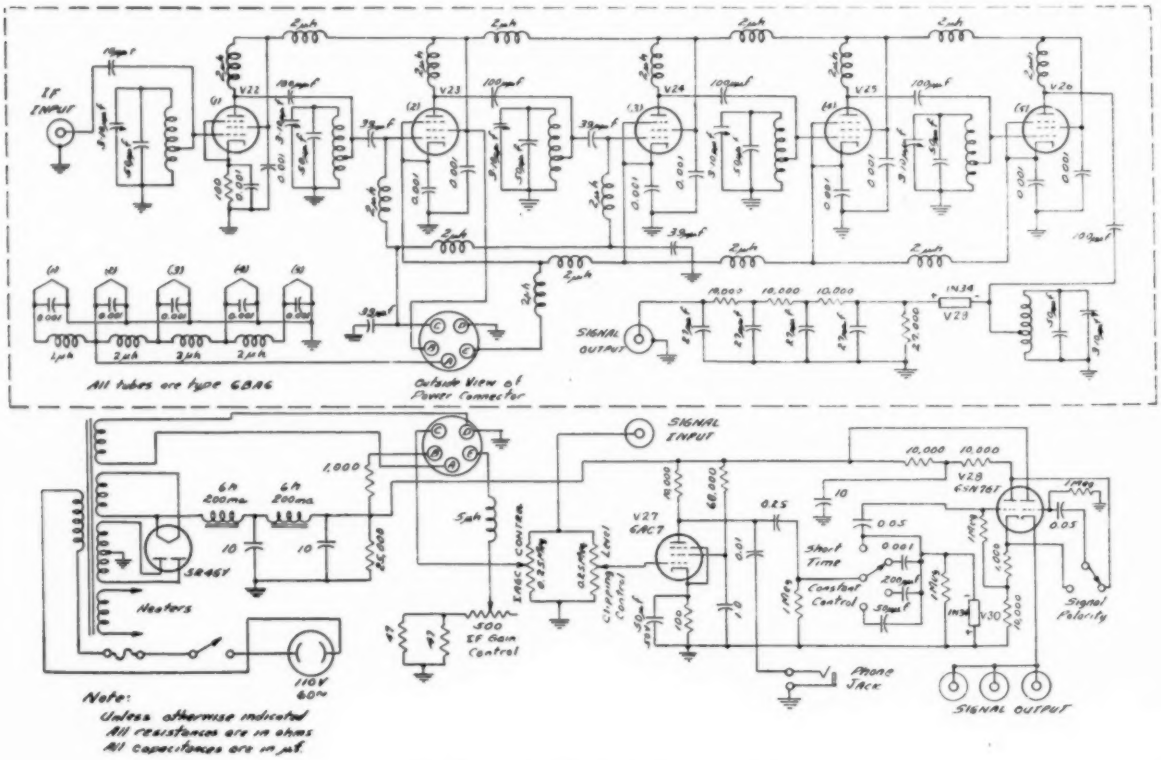



FIG. 9. Schematic diagram of the receiver.

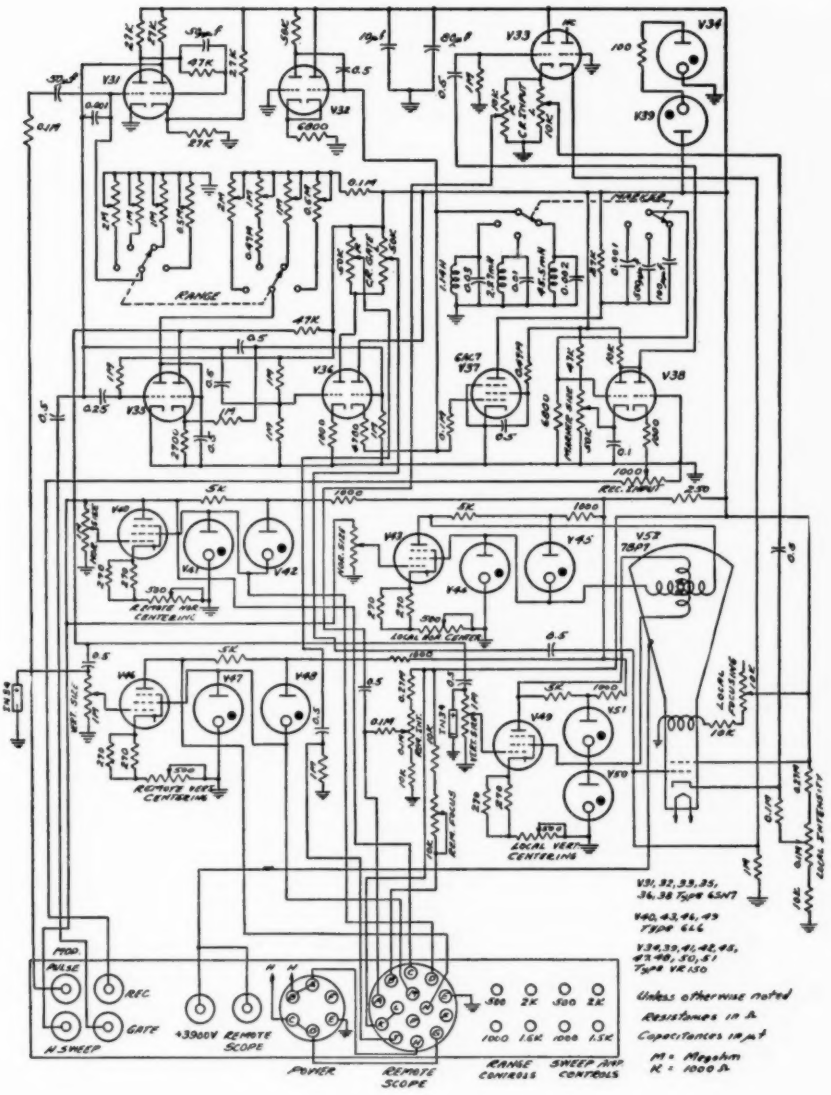



FIG. 10. Schematic diagram of the indicator unit.

are obtained by taking time exposures of the entire frequency sweep, which requires from 15 to 30 seconds. The camera used is an AAF type 05 35-mm recorder. Although this method of frequency sweep does not give a truly rectangular pattern because of the imperfection of the deflection system, it was chosen because it is easy to construct. A better pattern could have been obtained by moving the film past a fixed baseline; however, it was decided that the more complicated film drive was not warranted. The visual indicator is similar; however, it uses a large tube, a 7BP7, V52. The persistence of the pattern on the screen of this tube is sufficient to provide a continuous image when the equipment is in operation.

Four different values of full-scale height (range) are provided; 500, 1000, 1500, and 2000 kilometers. As an aid to scaling the records, both height and frequency markers are inserted electronically. The height markers, which appear as solid horizontal lines, are provided at intervals of 10, 50, or 200 kilometers. The frequency markers, which appear as dashed vertical lines, occur at intervals of 1 megacycle.

Figure 10 is a schematic diagram of the indicator unit. This includes all of the circuits for both indicators with the exception of the

photographic tube and its focusing and deflection coils.

The functions of this unit are initiated by a pulse from the modulator. This pulse actuates a trigger circuit, $V_{31}$, similar to the one in the modulator. The duration and timing of the sweep period are determined by the trigger circuit, which

(a) actuates the sweep generator,
(b) provides intensifying pulses for the cathode-ray tubes,
(c) actuates the height marker generator.

The sweep generator consists of a dual triode, $V_{35}$, connected so that the condenser in the plate circuit of the left-hand section is normally kept discharged. During the active part of the sweep interval the condenser is allowed to charge through a high resistance, producing a linear saw-tooth voltage. This voltage is rendered trapezoidal, as required for linear magnetic deflection, by a small amount of feedback from the trigger circuit, and is amplified and used for the vertical (height) sweep.

The intensifying pulses, which are made rectangular by means of a limiting amplifier, are applied to the control grids of the cathode-ray tube to increase the brightness of the pattern during the active sweep interval.

The height marker generator consists of a control tube, $V_{36}$, oscillator, $V_{32}$, and squaring stage, $V_{37}$. The control tube functions as a cathode follower, effectively placing a low impedance across the tuned circuit connected in its cathode. When the cathode current is suddenly interrupted by the trigger circuit, an oscillation is initiated, and is maintained by the oscillator $V_{32}$. The resulting square-wave output of $V_{37}$ is differentiated and clipped by $V_{38}$, and the series of negative pulses so obtained is applied to the cathodes of the cathode ray tubes. The signal voltage from the receiver is also applied to the same points.

Thus during each vertical sweep the signals and height marker are seen as bright dots. The dots are swept out into lines, which appear continuous, by the horizontal deflection circuits.

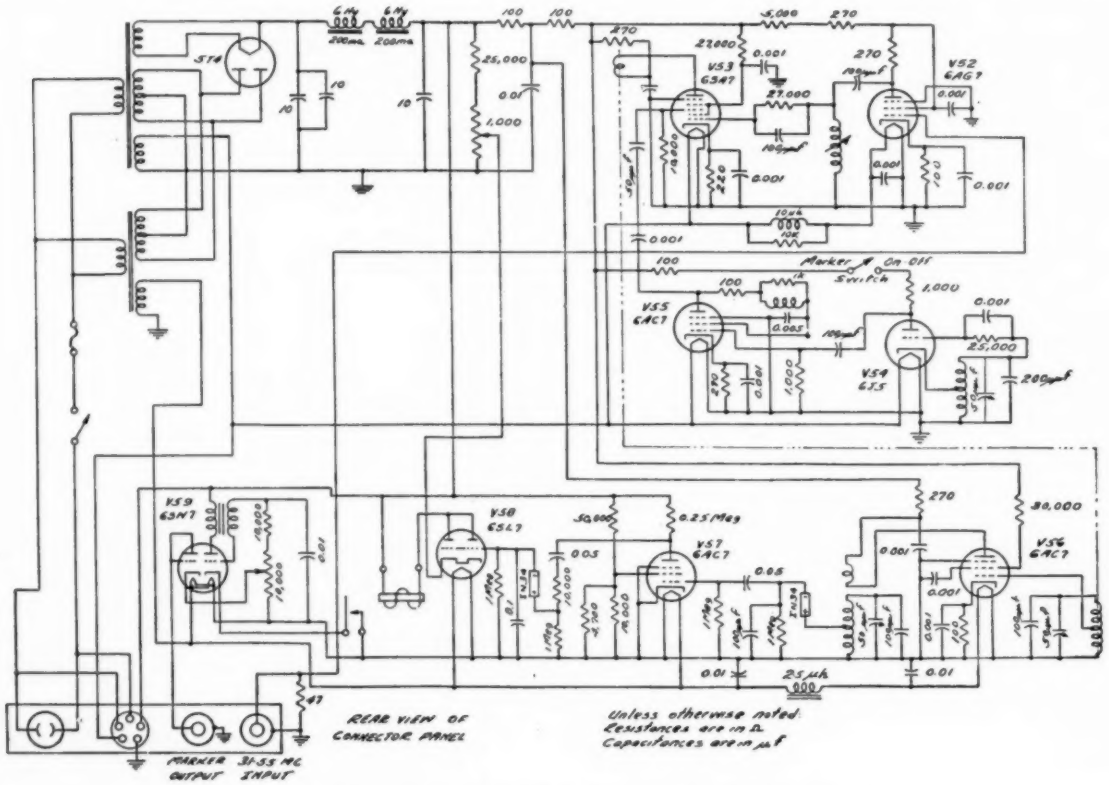



FIG. 11. Schematic diagram of the frequency marker generator.

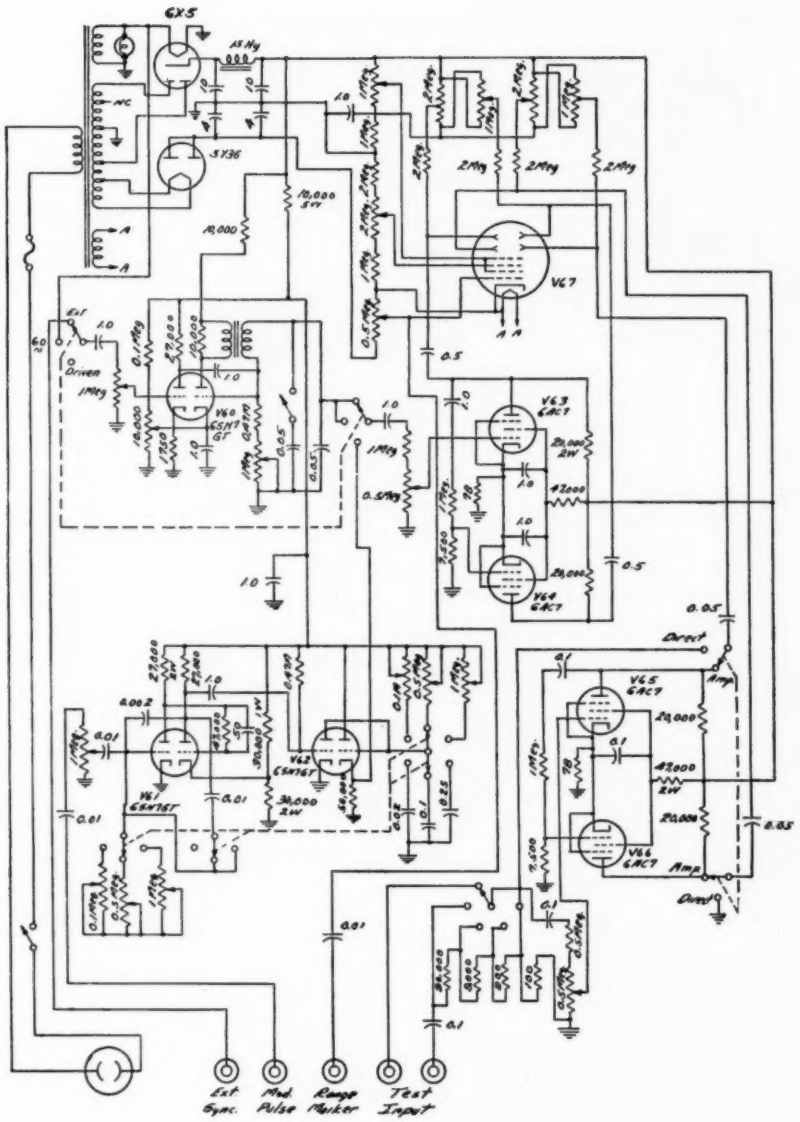



FIG. 12. Schematic diagram of the test oscilloscope.

Horizontal sweep voltage is obtained from a potentiometer mounted on the same shaft as the variable inductor used to obtain frequency variation.

The frequency marker generator, Fig. 11, consists of a standard frequency generator and a low sensitivity receiver. The receiver is tuned through the frequency range 1–25 megacycles in synchronism with the rest of the equipment. When its operating frequency coincides with one of the marker frequencies, which are spaced at intervals of 1 megacycle, a blocking oscillator is actuated, producing a dashed vertical line on the record.

The standard frequencies are produced by the 1-megacycle oscillator, $V_{54}$, and harmonic generator, $V_{55}$. Output from the common oscillator is amplified by $V_{52}$, and applied to the mixer, $V_{53}$, which also receives the standard frequencies. The mixer drives a 30-megacycle amplifier, $V_{56}$, a crystal-diode detector, and amplifiers $V_{57}$ and $V_{58}$. The output of the detector is a symmetrical pulse whose maximum occurs at the time it is desired to have a marker. By differentiating twice, a short pulse timed very nearly at the maximum is obtained. The short pulse is used to actuate a relay, which controls $V_{59}$, a blocking

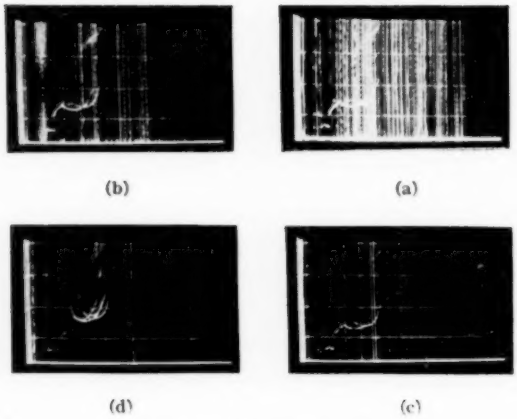



FIG. 13. Sample records.

oscillator. The blocking oscillator output produces the frequency marker.

The test oscilloscope, Fig. 12, is used for aligning the transmitter and receiver, and making any other necessary measurements in the equipment. It has provisions for both continuous and separately triggered sweeps. The continuous sweep is provided by the blocking oscillator $V_{60}$, while the triggered sweep is developed by the trigger circuit $V_{61}$ and sweep generator $V_{62}$. These circuits are similar to those described above, and will not be discussed here.

The test oscilloscope is normally connected to the receiver output, to be used for tuning the transmitter. However, a high voltage attenuator is provided so that pulses, either *DC* or *RF*, can be examined in any part of the receiver or transmitter.

## RESULTS

Figure 13 is a photograph of four records obtained with the equipment. These were made on June 15, 1948, at approximately 10:00 A.M., EDT. These records are as follows:

(a) Reflections from $E$, $F_1$, and $F_2$ layers; IAGC and differentiating circuits not connected. Interference from stations appears as bright vertical lines.

(b) Same as (a) with differentiation.

(c) Same as (a) with both differentiation and IAGC. Comparing this record with the one immediately above it shows that interference by all but the most powerful stations has been eliminated.

(d) An unusual record obtained shortly afterward showing complicated "splitting" in the $F_2$ region.

## ACKNOWLEDGMENT

The work reported in this paper was done under a contract with the Watson Laboratories, Air Material Command, Red Bank, New Jersey. The assistance of Mr. H. T. Lee and Mr. R. M. Hussey in the construction and assembly of this device is gratefully acknowledged.

# Work Functions and Conductivity of Oxide-Coated Cathodes*

G. W. MAHLMAN**
*Research Laboratory of Electronics,*** *Massachusetts Institute of Technology, Cambridge, Massachusetts*
(Received August 16, 1948)

The thermionic and photoelectric work functions and the temperature variation of coating conductivity have been measured for an oxide cathode. Thermionic and photoelectric currents were measured in both retarding and accelerating fields. The two work functions are found to be different, but not by the amount predicted by applying simple semi-conductor theory to the data. Photoelectric currents fit a Fowler plot rather well over a limited range of frequencies. The large decrease in work function with applied field, the poor "saturation" of thermionic currents typical of oxide cathodes, and the behavior of the photoelectric currents in accelerating fields all suggest that one is dealing with a "patchy" surface. Applying "checkerboard" patch theory to the experimental data, one finds that most of the data is accounted for by assuming patches about $3\times10^{-4}$ cm on a side differing in work function by about 0.2 volt.

## INTRODUCTION

IN 1940–1942, investigations were made independently by Nishibori, Kawamura, and Hirano[1] in Japan, and by Brown[2] at M.I.T. of some of the properties of the thermionic and photoelectric emission from oxide-coated cathodes. Nishibori, Kawamura, and Hirano measured the thermionic work function $\phi_{th}$, the photoelectric work function $\phi_{P.E.}$, and the temperature variation of coating conductivity $\sigma$. They claimed their results verified simple semi-conductor theory applied to the oxide cathode, but unfortunately the measurements of the two work functions were made on different cathodes. The variability observed in the important theoretical quantities $\phi_{th}$, $\phi_{P.E.}$, and $\sigma$ for oxide-coated cathodes makes it necessary that these quantities be determined for one and the same cathode.

Brown determined $\phi_{P.E.}$ from Fowler plots, since he found that the data fitted Fowler plots quite accurately near the threshold frequency. He obtained approximate equality between $\phi_{th}$ and $\phi_{P.E.}$, whereas Nishibori, Kawamura, and Hirano, using a different procedure to determine $\phi_{P.E.}$, did not obtain equality.

The work reported here is a continuation of the researches mentioned above. More complete details are available in the M.I.T. doctoral theses of the author and of Brown,[2] and in Technical Report No. 67 of M.I.T. Research Laboratory of Electronics.

By measurements of $\phi_{th}$, $\phi_{P.E.}$, and the temperature variation of $\sigma$, one should be able to learn something concerning energy differences between occupied and unoccupied energy levels in the oxide-coated cathode. Simple semi-conductor theory for an "excess" electronic semi-conductor yields the following formulas:[3,4]

$$J=10^{-6}Dn_b^{1/2}T^{5/4}\times\exp[\epsilon(-V-\tfrac{1}{2}E)/kT]\ \text{amp./cm}^2 \quad (1)$$

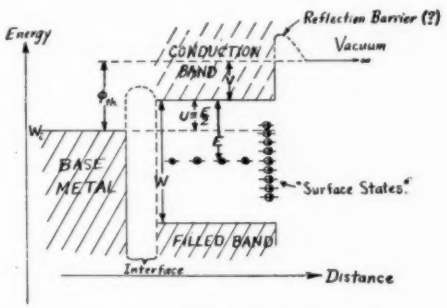



FIG. 1. Model of the oxide cathode.

* This is an abridged treatment of research submitted to the Department of Physics of Massachusetts Institute of Technology in partial fulfilment of the requirements for degree of Doctor of Science.

** Now at the Research Laboratory of the Linde Air Products Co., Union Carbide and Carbon Corp.

*** This work has been supported in part by the Signal Corps, the Air Materiel Command, and O.N.R.

[1] E. Nishibori, H. Kawamura, and K. Hirano, Proc. Phys. Math. Soc. (Japan) **22**, 378 (1940); *ibid.* **23**, 37 (1941).

[2] B. B. Brown, M.I.T. thesis, "An Investigation of Certain Electrical Properties of Oxide-Coated Cathodes," 1942 (unpublished).

[3] R. H. Fowler, *Statistical Mechanics* (Cambridge University Press, London, 1936).

[4] F. Seitz, *Modern Theory of Solids* (McGraw-Hill Book Company, Inc., New York, 1940).

TABLE I.

| ‡Cathode Temperature | Obs. Slope/Schottky Slope |
|---|---|
| 464°K | 17.85 at low fields<br>9.25 at high fields |
| 435°K | 16.10 at low fields<br>8.05 at high fields |
| 600°K | 15.50 at low fields<br>8.97 at high fields |
| 619°K | 14.70 at low fields |

so that

$$\phi_{th} = V + \tfrac{1}{2}E,$$

$$\phi_{P.E.} = V + E, \tag{2}$$

$$\sigma = \sigma_0 \exp(-\epsilon E/2kT) \tag{3}$$

where $\epsilon$ is the electronic charge, $D$ is the transmission coefficient, $n_b$ is the number of bound impurity atoms per cm³, $T$ is the absolute temperature, $k$ is Boltzman's constant, and energy gaps $E$ and $V$ are those shown in Fig. 1. If "surface states" exist, which are occupied by electrons, then one expects that $\phi_{P.E.}$ would not equal the energy gap $V+E$.

## EXPERIMENTAL PROCEDURE

The experimental tube consisted of a tantalum collector and two tantalum guard cylinders each ½ inch in diameter and 10 mm long, and a cathode containing a single embedded probe. Temperature of the cathode was measured by a nickel-tungsten thermocouple which measured base metal temperature, a method justified by the work of Fan.[5] Monochromatic illumination of the cathode was provided by light from a T-10 projection lamp after passing through a single monochromator. The effects of scattered light seemed to be inappreciable. Currents were measured down to about $10^{-14}$ ampere by a Victoreen VX-41 electrometer tube.

In an effort to avoid interface compounds which form due to impurities in the base metal, an electrolytic nickel was used as base metal. The work of Fineman and Eisenstein[6] and others indicated that interfaces form only very slowly on such a base metal. The presence of an appreciable thickness of interface compound could affect the measurement of $\sigma$, and consequently might lead to an erroneous value of the energy gap $E$ (Fig. 1).

A platinum probe wire ½ mil in diameter was used to determine $\sigma$. A coating of Radio Mixture C-51-2 was sprayed on the nickel core, the probe wire spiralled over the coated length, and another coating applied to hold the probe in place within the coating.

## THERMIONIC EMISSION

Measurements were made of thermionic currents in retarding fields. Results in brief were that the energy distribution of the thermions was Maxwellian for a temperature equal to that of the measured cathode temperature *only* for retarding fields of several tenths of a volt. Large deviations from the Maxwellian distribution occurred in the region of zero-field. This result disagrees with that reported by Fan.[5] Saturation of the thermionic currents in small accelerating fields was good, thus permitting determination of "zero-field" thermionic currents. As very careful measurements of thermionic currents in retarding fields are now being undertaken by C. S. Hung at M.I.T., the results obtained will not be presented in detail here.

Schottky plots of the thermionic currents in accelerating fields were straight at high fields (1500–5000 volts/cm) and low fields (up to about 300 volts/cm, and were curved at intermediate

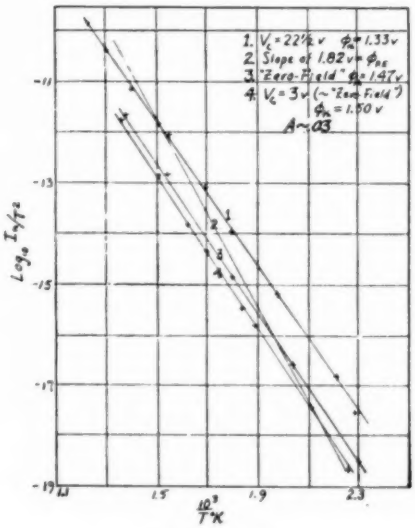



FIG. 2. Richardson plots.

[5] H. Y. Fan, J. App. Phys. **14**, 552 (1943).
[6] A. Fineman and A. S. Eisenstein, J. App. Phys. **17**, 633 (1946).

field strengths. The ratios of the observed slopes to the theoretical Schottky slopes are given in Table I. These ratios increase slightly with decreasing cathode temperature.

In Fig. 2 are shown Richardson plots. Plots 3 and 4 were taken at very nearly zero applied field, whereas plot 1 was taken at a collection voltage of $22\frac{1}{2}$ volts. Both $\phi_{th}$ and the Richardson "$A$" factor [in the Richardson Dushman emission equation $I=AT^2\exp(-\epsilon\phi/kT)$] were found to decrease as the applied field increased. The line numbered 2 in Fig. 4 is one having a slope corresponding to a work function of 1.82 volts, the zero-field $\phi_{P.E.}$. The measurements certainly do not indicate that $\phi_{th}=\phi_{P.E.}$.

## COATING CONDUCTIVITY MEASUREMENTS

If current is drawn through the oxide coating, the embedded probe assumes a certain potential with respect to the base metal or core of the cathode. As current through the coating increases, the probe becomes more and more positive with respect to the core in accordance with Ohm's law. The slopes of the experimental lines enable one to measure the resistance of a "slab" of oxide coating having a thickness equal to the separation of the probe from the core, from which $\sigma$ can be computed.

Another method of determining $\sigma$ is to consider the probe embedded in the coating as two essentially perfect conductors (the probe and the core) embedded in a medium of poor conductivity. If a potential is applied between probe and core, a current flows. The characteristics obtained are not Ohmic except near the origin and they show time changes. Difficulties due to polarization effects, the formation of dendrites, etc. are probably responsible for the observed behavior. The slope at the origin is taken as a measure of the resistance of the electrode configuration.

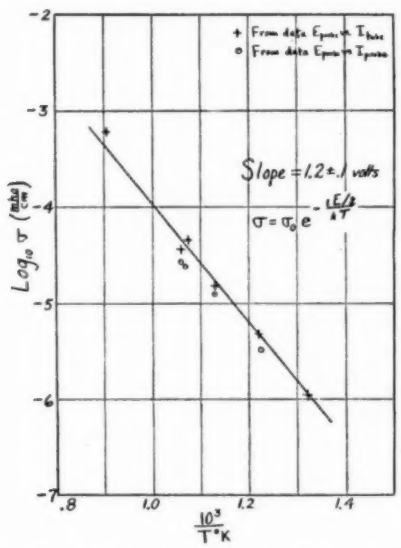



FIG. 3. Coating conductivity as a function of temperature.

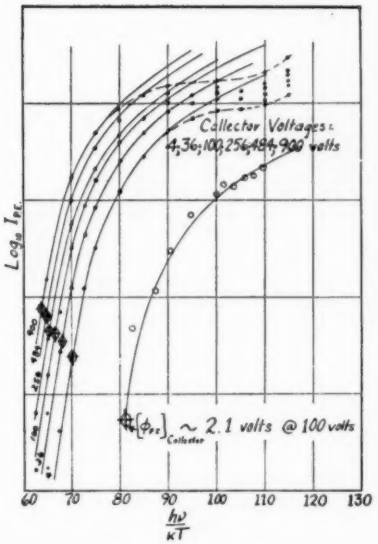



FIG. 4. Fowler plots.

Figure 3 is a plot of $\log\sigma$ *vs.* $1/T$ which according to Eq. (3) should have a slope $-\epsilon E/2k$. The two methods of measuring $\sigma$ give values in mho per cm which agree within experimental error. The value of $\frac{1}{2}E$ is determined to be $1.2\pm0.1$ volts.

## PHOTOELECTRIC EMISSION

The photoelectric currents from an oxide cathode fit Fowler plots quite accurately, a result first obtained by Brown.[2] In Fig. 4 are shown Fowler plots for different collection voltages. Also is a single, rather uncertain determination of the photoelectric work function of the collector anode.

Figure 5 shows the variation of $\phi_{P.E.}$ with applied field, using the data of Fig. 4. A similar variation in $\phi_{P.E.}$ with field was obtained by Brown.[2]

A rough check on the zero-field $\phi_{P.E.}$ was obtained by measuring photoelectric currents due to monochromatic radiation in retarding fields.

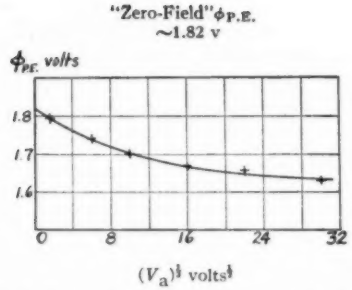



FIG. 5. Variation of photoelectric work function with applied voltage.

The values of $\phi_{P.E.}$ so determined seemed to agree better with the Fowler plot values than with values of $\phi_{P.E.}$ determined by the procedure employed by the Japanese workers Nishibori, Kawamura, and Hirano.[1]

Figure 6 shows the remarkable behavior of the photoelectric currents in accelerating fields, using light of different wave-lengths. Near the threshold ($\lambda = 6600$A), saturation is very poor. High energy photons ($\lambda = 4495$A) give good saturation, the slope of the experimental line being nearly equal to the theoretical Schottky slope. An intermediate wave-length ($\lambda = 5500$A) gives a curve which suggests the summation of two curves resembling those for $\lambda = 6600$A and $\lambda = 4495$A. Results qualitatively similar to these were obtained by Huxford,[7] who used red and green filters.

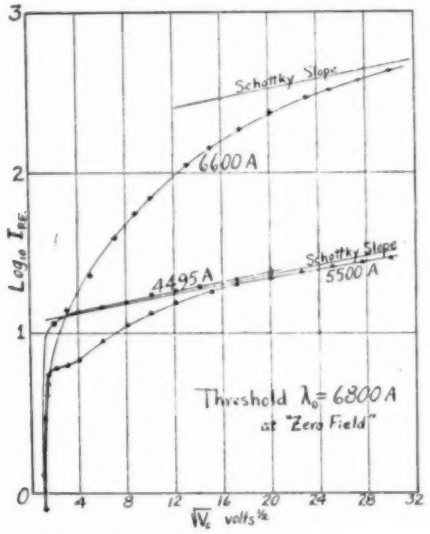



FIG. 6. "Schottky" plots for photoelectric currents at different wave-lengths.

[7] W. S. Huxford, Phys. Rev. **38**, 379 (1931).

## PATCH EFFECTS

It is believed much of the behavior observed, as represented for example by the data of Figs. 2, 5, and 6 is explainable in terms of "patch" effects. The theory of patches is discussed at length by Nottingham,[8] Becker,[9] Linford,[10] and others, and therefore will not be discussed in detail here.

The surface is visualized as consisting of many crystal faces, these faces being areas of slightly different work function. The potential barrier of a low work function area is lowered appreciably (order of tenths of volts) by application of only moderately strong fields, whereas the high work function barrier is lowered even less than for the "ideal" image barrier.

The large amount by which the low work function barrier is lowered as field is applied would account for the poor "saturation" of thermionic currents, and photoelectric currents near the threshold (the curve for $\lambda = 6600$A in Fig. 6), and also for the decreases observed in $\phi_{th}$ (Fig. 2) and $\phi_{P.E.}$ (Fig. 5). The behavior of photoelectric currents in accelerating fields (Fig. 6) is qualitatively completely explained by the assumption of a "patchy" emitting surface. The observed decrease in Richardson "$A$" factor with applied field results from a decrease in effective emitting area, since the patch fields act something like a grid, the applied field penetrating the field set up by the grid. The Schottky slopes observed would be expected to behave in the manner indicated by the data in Table I; that is, the ratio of observed slope to theoretical Schottky slope should decrease slightly with increasing temperature.

Figure 7 represents an attempt to estimate the approximate size of the patches, and the work function differences between them. The variation of work function $\phi$ with field $E$ is given by:

$$d\phi/dE = -Z_e \qquad (4)$$

where $Z_e$ is the "critical" distance from the emitting surface where the force pulling an electron toward the cathode equals the force of the applied field pulling it away from the cathode. Assuming smooth-cylinder geometry, the field $E$ acting on an electron is known as a function of

[8] W. B. Nottingham, Phys. Rev. **49**, 78 (1936).
[9] J. A. Becker, Rev. Mod. Phys. **7**, 95 (1935).
[10] L. B. Linford, Rev. Mod. Phys. **5**, 34 (1933).

distance $Z$ from the emitting surface, using Eq. (4).

The circles in Fig. 7 are computed points using the data in Fig. 5 to determine $d\phi_{P.E.}/dE$. Areas of the circles indicate estimated error, the dashed circle being a very uncertain point.

The $X$'s in Fig. 7 are computed points, using the data from a Schottky plot of thermionic currents taken at 535°K. The assumption was made that the increase in thermionic current with applied field was due entirely to decrease in $\phi_{th}$ with applied field, an assumption which is probably only approximately true. Data using other Schottky plots fall in the same region in Fig. 7.

The curved lines in Fig. 7 represent a first approximation to the low work function patch fields for the "checkerboard" patches proposed by Compton and Langmuir.[11] Square patches $b$ cm on a side are assumed, each patch differing in work function from its neighbors by the amount $\Delta V$. The field acting on an electron leaving the surface of a "checkerboard" emitter is then the sum of the patch field and the image field. Most of the data in Fig. 7 is accounted for by assuming patches about $3\times10^{-4}$ cm on a side, differing in work function by about 0.2 volt. The work of Nichols,[12] Eisenstein,[13] and others indicates that these values of $b$ and $\Delta V$ are quite reasonable.

A more complete discussion than is possible here of the theory involved in plotting the data in Fig. 7 is given by Linford,[10] whose analysis of data for other patchy emitters suggested treating experimental data in the manner shown in Fig. 7.

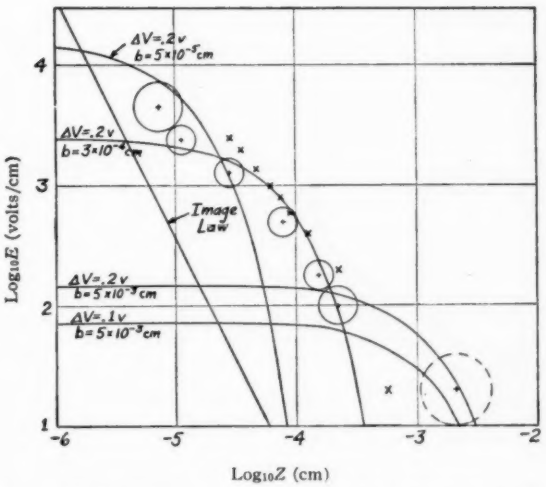



FIG. 7. Field acting on an electron as a function of distance from the emitter's surface.

## MODEL OF THE OXIDE CATHODE

The original purpose of this investigation was to determine $\phi_{th}$, $\phi_{P.E.}$, and the temperature variation of $\sigma$. Referring to Fig. 1, the energy gap $E$ was measured to about $2\times1.2=2.4$ volts, and $\phi_{th}\cong1.5$ volts (at zero field, of course).

If surface states played no appreciable part in photoelectric emission or if these states were nonexistent, we should have $\phi_{P.E.}=E+V$, assuming the photoelectrons came from the "bound" impurity levels. However, $\phi_{P.E.}$ was measured to be only about 1.8 volts, not $E+V=2.4+0.3=2.7$ volts ($V=\phi_{th}-U\cong1.5-1.2=0.3$ volt). Apparently then the photoelectrons came from occupied surface levels having higher energy than the impurity levels in the oxide interior.

One might expect the onset of a "volume photoelectric effect" at about 4600A [$\lambda A=hc/\epsilon(V+E)\cong12{,}400$ (2.7 volts)$=4600$A]. Such an effect would probably be small, as most of the photoelectrons would come from the surface, those in the interior losing energy in the lattice. Certainly no drastic increase in photoelectric current at 4600A was observed. There is no justification from the data obtained for the assumption that $\phi_{P.E.}=V+E$, as Nishibori, Kawamura, and Hirano[1] have claimed. For Eqs. (1)–(3) to be true, the experimental errors would have to be much greater than what they are estimated to be. This result is really not surprising when one considers that physically there is every reason to suppose that "surface states" do exist. These would arise not only from the interruption of lattice periodicity at the surface, giving rise to "Tamm" levels, but also from the probable existence of adsorbed barium on the oxide-coated cathode surface.

The photoelectric currents behave in many respects, though not in all respects, like those from metals. Thus the fit of the data to a Fowler plot is remarkably good near the threshold, and "Fermi tails" are observed. Certainly no marked "resonances" occur, as one could expect for electrons emitted from discrete energy levels.

[11] K. T. Compton and I. Langmuir, Rev. Mod. Phys. **2**, 123 (1930).
[12] M. H. Nichols, Phys. Rev. **57**, 297 (1940).
[13] A. S. Eisenstein, J. App. Phys. **17**, 434 (1946).

It should be pointed out that the measurements used to determine energy gaps are not beyond criticism. Thus it has been suggested that in measuring temperature variation of $\sigma$ by means of a probe, one is in fact measuring thermionic emission as it varies with temperature. This criticism seems a bit farfetched, however, as there is little in the observed behavior to suggest thermionic emission, and the measured values of $\frac{1}{2}E$ and $\phi_{th}$ differ by 0.3 volt.

The cathode on which the results of this study are based was not the most active attainable. It showed a zero-field value of $\phi_{th}=1.5$ volts and a value of 1.33 volts at a field of about 110 volts/cm. A more completely activated cathode would have had a value of $\phi_{th}$ of about 1.1 volts at a field of 110 volts/cm. The value of $E=2.4$ volts is equal to the highest value reported by Nishibori, Kawamura, and Hirano.[1] Their data indicates that a high $E$ goes with a high $\phi_{th}$. The $\phi_{P.E.}$ obtained here checks rather closely with the Nishibori *et al.* values, but is 0.3 to 0.7 volt higher than those of Huxford[7] or Brown.[2] Here again activity differences may account for some of the discrepancy.

The author wishes to express his thanks to Professor Wayne B. Nottingham and his group for many stimulating discussions concerning the results of this research. Thanks are due also to many members and technicians of the Research Laboratory of Electronics for assistance and advice in the construction of experimental tubes.

## A Positive-Replica Technique for Electron Microscopy*

C. M. Schwartz, A. E. Austin, and P. M. Weber
*Battelle Memorial Institute, Columbus, Ohio*

(Received August 18, 1948)

A positive-replica technique for electron microscopy has been developed, which reproduces the contour variations of the specimen surface, and permits direct visual interpretation of elevation. The method utilizes two resins, each mutually insoluble in the solvent for the other, specifically, polyvinyl alcohol plus Formvar. Shadow casting of the Formvar positive replica enhances detail and highlights the areas in relief in the original surface. Examples are given, in particular, applied to wear-test specimens.

### INTRODUCTION

In the course of an extended investigation** of journal-bearing behavior, now in progress, electron-microscopic studies of the effect of wear under load have found application in evaluating such phenomena as seizure, metal flow, and burnishing. The results are incomplete and will be described at a later date. However, certain requirements of specimen preparation, which were specific for this application, led to the development of a two-replica technique; a description thereof is warranted at this time by possible general utility.

The objective was the investigation of changes in surface contour during wear, in particular, at the first incidence of seizure under light loads. Although the bearings were run under conditions normally considered to be those of fluid-film lubrication, evidence of metal-to-metal contact at the high spots was sought.

The proposed test conditions imposed certain restrictions as to sample preparation, which did not appear to be met by existing techniques. The specimen itself, a $1\frac{1}{4}$-inch diameter cylindrical steel shaft, several inches long, provided the first restriction, namely, that the silica-polystyrene method[1] could not be used, owing to the impracticality of molding a sample of such shape. Furthermore, the examination had to be nondestructive, since it was planned to examine the same area of each shaft before and after a wear-

* Presented before the Annual Meeting of the Electron Microscope Society of America, the Franklin Institute, Philadelphia, December, 1947.

** ONR Contract No. N5-ORI-III, Task Order III.

[1] R. D. Heidenreich and V. G. Peck, J. App. Phys. **14**, 23 (1943).

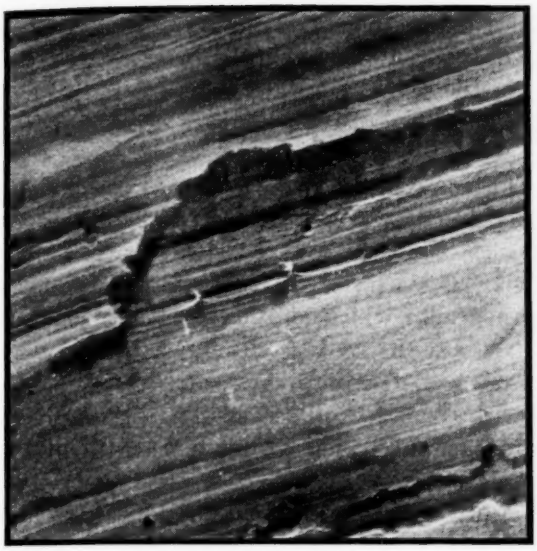


FIG. 1. Electron micrograph of negative replica of fine-ground steel shaft, shadow cast with gold, at 10:1 angle. 7000 ×.

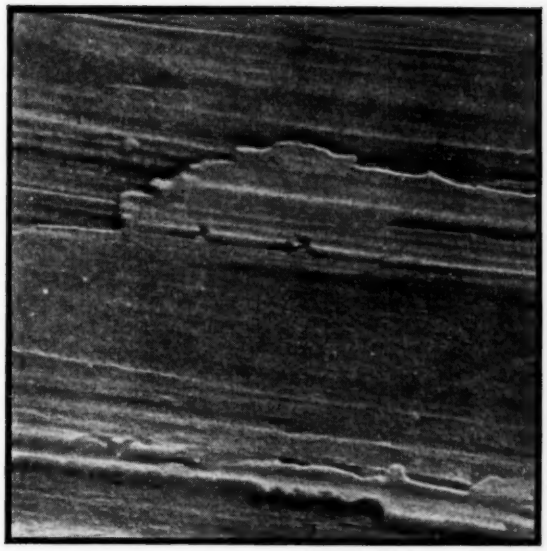


FIG. 2. Electron micrograph of positive replica of same field as in Fig. 1. Shadow cast with gold. 7000 ×.

test run. The principal interest was in recording details of the change in contour of the peaks and high spots of the metal surface during wear. The simple lacquer-film replica does not provide sufficient contrast and detail. These qualities can be accentuated by the method of shadow casting,[2] but when applied to a negative replica only the pits and valleys of the original surface are accentuated.

What is needed, therefore, is a two-step process yielding a positive replica, whose ridges, corresponding to high spots of the original surface, may be highlighted by shadowing. A solution of this problem has been found in the present work, in the selection of two plastic materials, each mutually insoluble in the solvent for the other.

## EXPERIMENTAL

The combination of water-soluble polyvinyl alcohol (PVA) and Formvar (soluble in ethylene dichloride) appears to satisfy the requirement of mutual insolubility in the respective solvents. PVA films disintegrate in the electron beam, so that the combination must be used with PVA as the intermediate negative replica.

It is necessary to obtain complete dissolution of the PVA film in water at one step in this method. Of the two types, *A* and *B*, offered by the manufacturer,[3] the former is the more soluble in cold water. This is supplied in several grades, varying in viscosity. Of those tried, the grade Elvanol A 51-05, which has the lowest viscosity, appeared to satisfy our requirements.

Attempts were first made to prepare thick negative replicas by impression of sheet PVA, using moderate heat and pressure, and with provision for subsequent chilling before release of the pressure. The results were not entirely satisfactory, owing mainly to reduced solubility of the material after this treatment. Films cast directly upon the metal surface, using a water latex of PVA, were then tried and found quite satisfactory. The possibility was considered that ferrous samples would rust in contact with the aqueous solution. However, a direct test, using electron diffraction, showed this effect to be inappreciable. A freshly abraded steel sample, successively filmed and stripped several times, showed no significant evidence of rust formation in subsequent examination of the metal surface by electron diffraction.

The detailed procedure follows: 2 drops of a 15 percent aqueous latex of PVA are spread over an area of sample of about 1 cm². The film is per-

[2] R. C. Williams and R. W. G. Wyckoff, J. App. Phys. **17**, 23 (1946).

[3] E. I. Du Pont de Nemours and Company, Electrochemicals Department.

FIG. 3. Electron micrograph of positive replica of fine-ground steel shaft, before running in bearing test. Shadow cast with gold. 7000 ×.

mitted to dry under a heat lamp, at a distance of several inches. Excessive drying is undesirable. When dry, the film is freed carefully at one edge, using a thin blade, whereupon it may be grasped with tweezers at the free edge and lifted completely from the specimen. A rigid but pliable negative replica results. The replica surface is coated with a layer of 0.75 percent to 1.0 percent Formvar in ethylene dichloride,[4] and allowed to dry quickly in a vertical position before the heat lamp, the excess solution being simultaneously drawn off the bottom edge with a piece of filter paper.

The composite film is conveniently cut into $\frac{1}{8}$-inch squares, which are placed carefully upon the surface of clean water, so as to float PVA side down. The latter dissolves, and the remaining Formvar film, which continues to float, is washed free of any PVA by continuous dilution with warm water. The Formvar positive film is picked up with the standard screen-disk specimen support, and subsequently handled in the usual manner.

An interesting and extremely useful corollary procedure has developed from the consideration of mutual insolubility of these two resins, and from the observed adhesion of PVA. In the preparation of Formvar negative replicas, the PVA latex has been found very satisfactory for stripping the Formvar film off metal and other surfaces. The method is somewhat similar to that which employs gelatine.[5] A PVA latex lens is applied over the dried Formvar film, the water allowed to dry as above, whereupon the composite film is peeled completely with no difficulty.

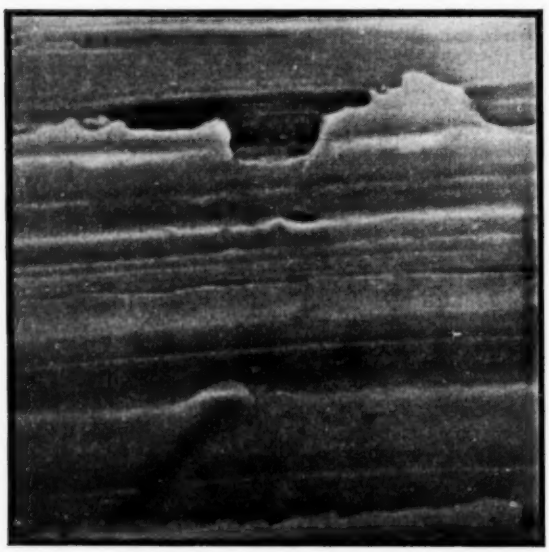


FIG. 4. Electron micrograph of positive replica of same field as in Fig. 3, after wear test. Gold shadowed. 7000 ×.

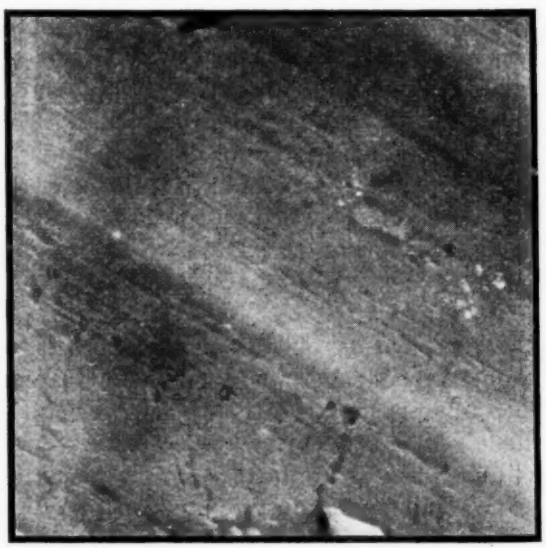


FIG. 5. Electron micrograph of positive replica of field adjacent to that of Figs. 3 and 4. Gold shadowed. 7000 ×.

[4] V. J. Schaeffer, J. App. Phys. **13**, 427 (1942).
[5] V. J. Schaeffer, Phys. Rev. **62**, 495 (1942).

The entire area of Formvar under the PVA comes free. The composite film is separated by dissolution of the PVA as described above. The advantages of this method over the Scotch-Tape method[5] are apparent in the complete absence of strain in the replica, and in the certainty of complete stripping. The gelatine method suffers in comparison, owing to the much greater adhesion of the PVA, and the simplicity of manipulation thereof.

## EXAMPLES OF ELECTRON MICROGRAPHS OF POSITIVE REPLICAS

This positive-replica technique was initially developed for the purpose of furnishing a representation of the surface contours of journal bearings and corresponding shafts. Figures 1 and 2 exemplify the advantages of this technique in providing a contour representation which is readily interpreted by inspection. These micrographs are, respectively, from negative and positive replicas of the same area of a cylindrically ground steel shaft, before the shaft was run in a bearing. The surface finish consisted of a normal commercial fine grind, giving a profilometer reading of about 5 microinches r.m.s. Figure 1, from a gold-shadowed negative Formvar replica, yields a false impression of elevation. With experience, of course, the contour may be correctly interpreted; however, the positive replica, Fig. 2, is correctly interpreted without conscious effort on the part of the observer. The presence of peened-over burrs or fins in Fig. 2 is worthy of note.

Figures 3 and 4 are positive replicas of the same area of the shaft before and after wear, respectively. The wear test run was made under light load conditions, 500 p.s.i. for 20 minutes against an aluminum bearing, in order to avoid completely "smearing out" all areas photographed prior to this test. The field of Fig. 4 has obviously been damaged only moderately. The general contour of the peened-over region is recognizable as corresponding to that in Fig. 3, although a portion of fin has been chipped out.

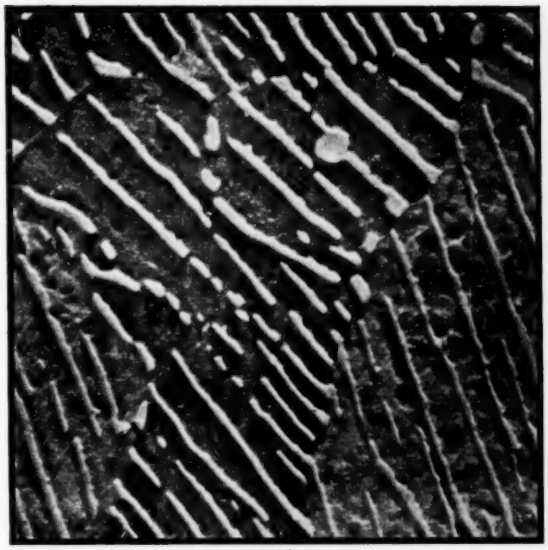


FIG. 6. Electron micrograph of gold-shadowed positive replica of pearlitic field from eutectoid steel. Specimen mechanically polished, etched in HCl—$FeCl_3$-Methanol. 7000 ×.

Figure 5 is a replica of a field adjacent to that of Fig. 4, about 50 microns distant. The original contours are completely obliterated; this type of topography did not exist on the original ground shaft. This field was obviously an area of high-pressure contact during the run, resulting in a well-burnished surface.

The positive-replica technique has been applied to a few types of metallographic samples, in an exploratory manner. Figure 6 illustrates the rather excellent possibilities in this field of application. The specimen was a eutectoid steel, austenitized and isothermally transformed to yield coarse pearlite. The indications of preferential etch attack, with respect to grain orientation, are quite apparent.

# A Mathematical Theory of Photo-Viscoelasticity

RAYMOND D. MINDLIN
*Department of Civil Engineering, Columbia University, New York, New York* and *Bell Telephone Laboratories, Murray Hill, New Jersey*

(Received August 23, 1948)

A phenomenological law of optical birefringence, induced by small strain in viscoelastic materials, is formulated and applied to an idealized medium. The resulting optical-stress-strain-time-temperature relations are used in finding conditions under which models made of viscoelastic materials may be employed in the photoelastic method of solving boundary value problems in the linear theory of elasticity.

## I. INTRODUCTION

A DESCRIPTION of photoelastic phenomena in viscoelastic media must include time-dependent relations between three second-rank tensors: birefringence, stress, and strain. From the viewpoint of mechanics of continua it is desirable to include, also, space-dependence (non-homogeneous strain) and certain aspects of temperature dependence. All of these variables can be incorporated in a relatively simple theory if the system is assumed to be linear. The attribute of linearity is not attained, of course, without some loss of generality. The chief sacrifices are finite strain and non-linear stress-strain and stress-strain-rate relations. In addition, thermodynamic effects have been neglected. What remains, however, is of considerable interest and wide applicability.

In Section II, a strain-optical law is formulated for small strain in viscoelastic media. The stress-strain properties of the medium to which the law is applied are described in Section III. These stress-strain properties are both more and less general than those usually employed for similar purposes. They are more general in including instantaneous elastic compressibility and time-dependent viscosity and elasticity, both of which are considered to be of interest in photo-viscoelastic phenomena. The medium is less general in including only one element, each, of instantaneous elasticity, retarded elasticity and flow. However, as indicated at appropriate places in the text, most of the results take the same form regardless of the number of shear moduli, viscosity coefficients, and time derivatives entering into the stress-strain relations. The general strain-optical law is applied to the idealized medium in Section IV, and conclusions are drawn regarding the relations between birefringence and stress and strain in various types of viscoelastic media.

The next two sections deal with the relations between stress, strain, and the optical properties that are most easily measured, and it is shown that, in general, the relations are not simple. This is followed by a study of the conditions under which these relations reduce to forms similar to those for an elastic medium.

Sections VIII to XII contain an investigation of the conditions under which the birefringence in a viscoelastic body may be translated directly to the stress in an elastic body of the same shape. The results are summarized in Section XIII. The essence of the conclusions is that viscoelastic materials may be used for models in photoelasticity except if there are loads which move over the surface of the model.

## II. GENERAL STRAIN-OPTICAL EQUATION

Consider a continuous, viscoelastic medium of which the mechanical model[1] is an assemblage of $m$ elastic elements and an equal, greater, or lesser number of viscous elements. Let $\Phi_k'$, $k=1, 2, \cdots m$, be the small, pure-strain dyadic[2] of the $k$th elastic element and let $\theta_k'$ be its mean stretch (i.e., one-third of its cubical dilatation). The strain deviation of the $k$th elastic element

[1] See T. Alfrey, *Mechanical Behavior of High Polymers* (Interscience Publishers, Inc., New York, 1948) as a general reference on the use of mechanical models of viscoelastic materials and especially Appendices II and III in which are summarized Alfrey's investigations of homogeneous and non-homogeneous stress and strain in a very general viscoelastic medium.

[2] For the dyadic notation used in this paper, see C. E. Weatherburn, *Advanced Vector Analysis* (G. Bell and Sons, London, 1928).

will be represented by $\Phi_{dk}' = \Phi_k' - \theta_k' \mathbf{I}$ where $\mathbf{I}$ is the idemfactor.

Let $\mathbf{\Pi}$ be the index dyadic, so that

$$\mathbf{r} \cdot \mathbf{\Pi} \cdot \mathbf{r} = 1 \qquad (2.01)$$

is the equation of the index (Fresnel) ellipsoid[3] if $\mathbf{r}$ is the radius vector drawn from the center of the ellipsoid to its surface.

If it is assumed that the birefringence arises solely from the deformation of the elastic elements, the most general linear, isotropic relation that can exist between birefringence and strain is

$$\mathbf{\Pi} - n_0^{-2}\mathbf{I} = \sum_{k=1}^{m} (c_k \Phi_{dk}' + d_k \theta_k' \mathbf{I}), \qquad (2.02)$$

where $c_k$ and $d_k$ are strain-optical coefficients and $n_0$ is the index of refraction of the unstrained medium.

Before the birefringence can be expressed in terms of the macroscopic strain and stress, it is necessary to specify the mechanical model of the viscoelastic medium and its stress-strain properties.

## III. STRESS-STRAIN RELATIONS FOR AN IDEALIZED VISCOELASTIC MATERIAL

Aside from failure, the important mechanical phenomena exhibited by viscoelastic materials are *instantaneous elasticity*, *retarded elasticity*, and *flow*. All three of these are exhibited, in turn, by a simple mechanical model, illustrated in Fig. 1, which has been used frequently to represent an idealized viscoelastic medium.[1]

In Fig. 1, the spring marked (1)′ represents the instantaneous elasticity. This element is assumed to possess linear, isotropic elasticity, with shear modulus $G_1$ and Poisson's ratio $\nu_1$. It is customary to assume that viscoelastic materials are incompressible and this is, indeed, very nearly true especially for retarded elasticity and flow. However, if complete incompressibility is assumed, the mean strain-optical coefficient $d_k$ vanishes. Since there is not sufficient experimental evidence to warrant the omission of $d_k$, the attribute of instantaneous elastic compressibility will be retained until non-homogeneous stress is considered, at which time (Section IX) the reason for the restriction to $\nu_1 = \frac{1}{2}$ will become apparent.

The spring marked (2)′ and the dash-pot marked (2)″ in Fig. 1, together represent the retarded elasticity. Incompressibility is assumed for this phenomenon and, hence, the linear, isotropic, elastic element (2)′ has only a shear modulus $G_2$ and the linear, isotropic viscous element (2)″ has only a viscosity coefficient $\eta_2$. Finally, the dash-pot marked (3)″ represents the incompressible flow with viscosity coefficient $\eta_3$.

In some materials $G_1$ and $G_2$ are quite insensitive to temperature over a limited range, whereas $\eta_2$ and $\eta_3$ depend strongly on temperature. Use of this attribute is made in the freezing method of three-dimensional photoelasticity. In order to include a wide variety of materials, $G_1$ and $G_2$ will also be considered temperature dependent. It will be assumed that the temperature is a known function of time and independent of the space coordinates and hence we shall consider that *the viscosity coefficients and shear moduli are functions of time only*.

The following notation will be used:

$\Phi_1'$, $\Phi_2'$, $\Phi_2''$, $\Phi_3''$ = pure strain dyadic of elements (1)′, (2)′, (2)″, (3)″, respectively,

$\theta_1'$, $\theta_2'$, $\theta_2''$, $\theta_3''$ = mean stretch of elements (1)′, (2)′, (2)″, (3)″, respectively,

$\Psi_1'$, $\Psi_2'$, $\Psi_2''$, $\Psi_3''$ = stress dyadic of elements (1)′, (2)′, (2)″, (3)″, respectively.

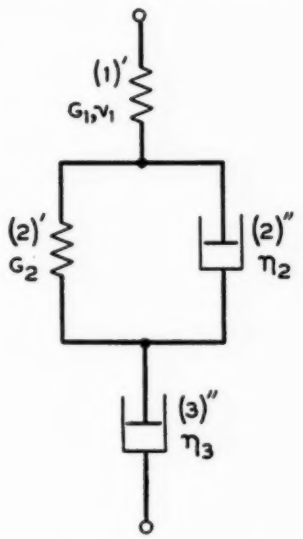



FIG. 1. Mechanical (spring and dash-pot) model of idealized viscoelastic medium.

[3] E. G. Coker and L. N. G. Filon, *A Treatise on Photo-Elasticity* (Cambridge University Press, London, 1932), p. 17.

$\Theta_1'$, $\Theta_2'$, $\Theta_2''$, $\Theta_3''$ = mean tension of elements (1)′, (2)′, (2)″, (3)″, respectively,

$\mathbf{\Phi}$ = macroscopic pure strain dyadic,
$\theta$ = macroscopic mean stretch,
$\mathbf{\Phi}_d = \mathbf{\Phi} - \theta\mathbf{I}$ = macroscopic strain deviation,
$\mathbf{\Psi}$ = macroscopic stress dyadic,
$\Theta$ = macroscopic mean tension,
$\mathbf{\Psi}_d = \mathbf{\Psi} - \Theta\mathbf{I}$ = macroscopic stress deviation.

The assumption of incompressibility of retarded elasticity and flow requires

$$\theta_2' = \theta_2'' = \theta_3'' = 0, \tag{3.01}$$

i.e., the pure strain dyadics $\mathbf{\Phi}_2'$, $\mathbf{\Phi}_2''$, $\mathbf{\Phi}_3''$ are equal to their respective deviations. The geometry of the model requires

$$\mathbf{\Phi}_2' = \mathbf{\Phi}_2'', \tag{3.02}$$

$$\mathbf{\Phi}_1' + \mathbf{\Phi}_2' + \mathbf{\Phi}_3'' = \mathbf{\Phi}. \tag{3.03}$$

The geometry and equilibrium of the model require

$$\mathbf{\Psi}_1' = \mathbf{\Psi}_2' + \mathbf{\Psi}_2'' = \mathbf{\Psi}_3'' = \mathbf{\Psi}, \tag{3.04}$$

$$\Theta_1' = \Theta_2' + \Theta_2'' = \Theta_3'' = \Theta. \tag{3.05}$$

Finally, (3.01) and (3.03) require

$$\theta_1' = \theta. \tag{3.06}$$

Then, neglecting thermodynamic effects, the stress-strain relations are, for element (1)′,

$$\mathbf{\Psi}_d = 2G_1(\mathbf{\Phi}_1' - \theta\mathbf{I}), \quad = 2G_1\mathbf{\Phi}_{d1}', \tag{3.07}$$

$$(1-2\nu_1)\Theta = 2G_1(1+\nu_1)\theta; \tag{3.08}$$

for elements (2)′ and (2)″,

$$\mathbf{\Psi}_d = 2G_2\mathbf{\Phi}_2' + 2\eta_2\frac{\partial\mathbf{\Phi}_2'}{\partial t}; \tag{3.09}$$

and, for element (3)″,

$$\mathbf{\Psi}_d = 2\eta_3\frac{\partial\mathbf{\Phi}_3''}{\partial t}, \tag{3.10}$$

where $t$ is the time.

Equations (3.07), (3.09), and (3.10) may be solved for $\mathbf{\Phi}_{d1}'$ and $\mathbf{\Phi}_{d2}'$ (which are required for the strain-optical equation) by using (3.03) to eliminate $\mathbf{\Phi}_3''$ and obtaining the additional equation, that is required, by differentiating (3.07) with respect to $t$. The results are

$$2G_1\mathbf{\Phi}_{d1}' = \mathbf{\Psi}_d, \tag{3.11}$$

$$2G_2\mathbf{\Phi}_{d2}' = \left(1 + \frac{\eta_2}{\eta_3} + \eta_2\frac{dG_1^{-1}}{dt} + \frac{\eta_2}{G_1}\frac{\partial}{\partial t}\right)\mathbf{\Psi}_d - 2\eta_2\frac{\partial\mathbf{\Phi}_d}{\partial t}, \tag{3.12}$$

in which $d/dt = (dT/dt)(d/dT)$, where $T$ is the temperature.

The relation between $\mathbf{\Psi}_d$ and $\mathbf{\Phi}_d$, independent of $\mathbf{\Phi}_1'$, $\mathbf{\Phi}_2'$, $\mathbf{\Phi}_3''$, is obtained by substituting (3.12) into (3.09), with the result

$$\left[\frac{G_2}{\eta_3} + F(t) + \left(1 + \frac{G_2}{G_1} + \frac{\eta_2}{G_1}\frac{dG_1^{-1}}{dt} + \frac{G_2}{G_1}\frac{d(\eta_2/G_2)}{dt} + \frac{\eta_2}{\eta_3}\right)\frac{\partial}{\partial t} + \frac{\eta_2}{G_1}\frac{\partial^2}{\partial t^2}\right]\mathbf{\Psi}_d = 2\left[G_2\left(1 + \frac{d(\eta_2/G_2)}{dt}\right)\frac{\partial}{\partial t} + \eta_2\frac{\partial^2}{\partial t^2}\right]\mathbf{\Phi}_d, \tag{3.13}$$

where

$$F(t) = \frac{d(\eta_2/\eta_3)}{dt} + \frac{G_2(1+\eta_2)}{G_1\eta_3}\frac{d(G_1/G_2)}{dt} + \left(G_1 + G_2 + \frac{G_1\eta_2}{\eta_3} + G_2\frac{d(\eta_2/G_1)}{dt}\right)\frac{dG_1^{-1}}{dt} + \eta_2\frac{d^2G_1^{-1}}{dt}.$$

*Equations (3.08) and (3.13) describe the relation between stress and strain in the idealized viscoelastic medium represented by Fig. 1.*

Let

$$\mathbf{P}[\ ] = \left[\frac{G_2}{\eta_3} + F(t) + \left(1 + \frac{G_2}{G_1} + \frac{\eta_2}{G_1}\frac{dG_1^{-1}}{dt} + \frac{G_2}{G_1}\frac{d(\eta_2/G_2)}{dt} + \frac{\eta_2}{\eta_3}\right)\frac{\partial}{\partial t} + \frac{\eta_2}{G_1}\frac{\partial^2}{\partial t^2}\right], \tag{3.14}$$

$$\mathbf{Q}[\ ] = \left[G_2\left(1 + \frac{d(\eta_2/G_2)}{dt}\right)\frac{\partial}{\partial t} + \eta_2\frac{\partial^2}{\partial t^2}\right]. \tag{3.15}$$

Then (3.13) becomes[4]

$$\mathbf{P}[\mathbf{\Psi}_d] = 2\mathbf{Q}[\mathbf{\Phi}_d]. \tag{3.16}$$

One of the materials used very frequently in photoelasticity is a cross-linked polymer (Bakelite BT-61-893) which exhibits instantaneous and retarded elasticity, but no flow. For materials of this type $\eta_3 = \infty$, and $G_1$ and $G_2$ are constants.

[4] See reference 1, p. 561, for comparison.

Hence, (3.13) reduces to

$$\left[1+\frac{G_2}{G_1}+\frac{1}{G_1}\frac{d\eta_2}{dt}+\frac{\eta_2}{G_1}\frac{\partial}{\partial t}\right]\Psi_d = 2\left[G_2+\frac{d\eta_2}{dt}+\eta_2\frac{\partial}{\partial t}\right]\Phi_d. \quad (3.17)$$

Setting $\eta_3=\infty$, $G_1$=constant, in (3.12), and using (3.17), the expression for $\Phi_{d2}'$ reduces to

$$2G_1\Phi_{d2}'=\left(1+\frac{1}{G_2}\frac{d\eta_2}{dt}\right)(2G_1\Phi_d-\Psi_d). \quad (3.18)$$

## IV. STRESS-STRAIN-OPTICAL EQUATION FOR THE IDEALIZED MEDIUM

The general strain-optical equation (2.02) reduces, for the idealized medium, to

$$\mathbf{\Pi}-n_0^{-2}\mathbf{I}=c_1\Phi_{d1}'+c_2\Phi_{d2}'+d_1\theta\mathbf{I}. \quad (4.01)$$

Thus three strain-optical coefficients, $c_1$, $c_2$, $d_1$, are required to describe the optical properties of the idealized, isotropic, viscoelastic medium as opposed to two coefficients for an isotropic, elastic medium. Two of the coefficients in (4.01), namely, $c_1$ and $c_2$, are *relative* strain-optical coefficients and contribute only to relative phase retardation, as will be shown below. The third coefficient, $d_1$, is a *mean* strain-optical coefficient whose magnitude can be determined only by measurement of absolute phase retardation. All of the coefficients may be temperature-dependent.

The relation between birefringence and macroscopic stress and strain is obtained by substituting, into (4.01), the expressions for $\Phi_{d1}'$ and $\Phi_{d2}'$ given in (3.11) and (3.12), with the result*

$$\mathbf{\Pi}-n_0^{-2}\mathbf{I}=\frac{1}{2}\left[\frac{c_1}{G_1}+\frac{c_2}{G_2}\left(1+\frac{\eta_2}{\eta_3}+\eta_2\frac{dG_1^{-1}}{dt}+\frac{\eta_2}{G_1}\frac{\partial}{\partial t}\right)\right]\Psi_d-\frac{c_2\eta_2}{G_2}\frac{\partial\Phi_d}{\partial t}+d_1\theta\mathbf{I}. \quad (4.02)$$

Let

$$\mathbf{R}[\ ]=\frac{1}{2}\left[\frac{c_1}{G_1}+\frac{c_2}{G_2}\left(1+\frac{\eta_2}{\eta_3}+\eta_2\frac{dG_1^{-1}}{dt}+\frac{\eta_2}{G_1}\frac{\partial}{\partial t}\right)\right], \quad (4.03)$$

$$\mathbf{S}[\ ]=\frac{c_2\eta_2}{2G_2}\frac{\partial}{\partial t}. \quad (4.04)$$

Then the stress-strain-optical equation becomes

$$\mathbf{\Pi}-n_0^{-2}\mathbf{I}=\mathbf{R}[\Psi_d]+2\mathbf{S}[\Phi_d]+d_1\theta\mathbf{I}. \quad (4.05)$$

In the case of the medium exhibiting no flow ($\eta_3=\infty$, $G_1$ and $G_2$=constants), the relation between birefringence and stress and strain is obtained by substituting (3.11) and (3.18) into (4.01), with the result

$$\mathbf{\Pi}-n_0^{-2}\mathbf{I}=\frac{1}{2G_1}\left[c_1-c_2\left(1+\frac{1}{G_2}\frac{d\eta_2}{dt}\right)\right]\Psi_d+c_2\left(1+\frac{1}{G_2}\frac{d\eta_2}{dt}\right)\Phi_d+d_1\theta\mathbf{I}. \quad (4.06)$$

It is interesting to observe that, in the medium exhibiting flow, the relative birefringence depends upon the stress, stress-rate, and strain-rate and is independent of the strain itself. In the medium exhibiting no flow, the relative birefringence is simply a linear function of the stress and strain and is independent of both the stress- and strain-rate. In the case of a medium exhibiting flow but no retarded elasticity, $\Phi_2'$ is absent from (4.01), and the expression for birefringence is obtained by setting $c_2=0$ in (4.02), whence the relative birefringence depends only on the stress.

Stress-strain-optical equations analogous to (4.02) can be obtained for a viscoelastic medium represented by a model having any number and arrangement of elastic and viscous elements since the system of stress-strain equations for the elements can always be solved for the $\Phi_{dk}'$ to obtain expressions corresponding to (3.11) and (3.12). A similar situation obtains in regard to stress-strain relations analogous to (3.13). Thus, wherever the linear operators **P**[ ], **Q**[ ], **R**[ ], **S**[ ] appear, they may be replaced by another set of linear operators appropriate to a different material.

## V. OPTICAL MEASUREMENTS

The optical quantities which are most easily, and therefore most commonly, measured are relative phase retardation ($\rho$) and amplitude ratio ($\gamma$).

In the case of homogeneous birefringence, the

* It may be observed that, inasmuch as the elastic elements are assumed to be linearly elastic, the general strain-optical relation, Eq. (2.02), might just as well have been expressed in terms of the stresses in the elastic elements, with similar results.

relative phase retardation is given by

$$\rho = 2\pi L(n_1 - n_2)/\lambda, \tag{5.01}$$

where $L$ is the path length, $\lambda$ is the wave-length in vacuum, and $n_1$ and $n_2$ are the principal indices of refraction in the plane of the wave front (secondary principal indices). The amplitude ratio gives, directly, the angle $\varphi$ which a polarizing axis makes with a reference direction in the plane of the wave front.

In the case of non-homogeneous birefringence, $n_1 - n_2$ and $\varphi$ may be obtained from measurements of $\rho$ and $\gamma$ by solving Neumann's equations:[5]

$$\frac{d\gamma}{dz} = -\frac{d\varphi}{dz}\cos\rho, \tag{5.02}$$

$$\frac{d\rho}{dz} - \frac{2\pi}{\lambda}(n_1 - n_2) = 2\frac{d\varphi}{dz}\cot 2\gamma \sin\rho. \tag{5.03}$$

We require formulas relating $n_1 - n_2$ and $\varphi$ to the stress and strain in a viscoelastic medium. Since $n_1$ and $n_2$ differ little from $n_0$ and from each other, $n_1 - n_2$ is nearly equal to $n_0^3(n_2^{-2} - n_1^{-2})/2$, and hence it will be sufficient to find a relation involving $n_2^{-2} - n_1^{-2}$.

## VI. RELATIVE STRESS-STRAIN-OPTICAL RELATIONS IN THE PLANE OF THE WAVE FRONT

In terms of the orthogonal unit vectors **i**, **j**, **k**, parallel to the axes of $x$, $y$, $z$, respectively, the nonian forms of the dyadics $\mathbf{\Pi}$, $\mathbf{\Phi}$, $\mathbf{\Psi}$ are

$$\mathbf{\Pi} = n_{xx}^{-2}\mathbf{ii} + n_{yy}^{-2}\mathbf{jj} + n_{zz}^{-2}\mathbf{kk} + n_{yz}^{-2}(\mathbf{jk}+\mathbf{kj}) + n_{zx}^{-2}(\mathbf{ki}+\mathbf{ik}) + n_{xy}^{-2}(\mathbf{ij}+\mathbf{ji}), \tag{6.01}$$

$$\mathbf{\Phi} = \epsilon_{xx}\mathbf{ii} + \epsilon_{yy}\mathbf{jj} + \epsilon_{zz}\mathbf{kk} + \epsilon_{yz}(\mathbf{jk}+\mathbf{kj}) + \epsilon_{zx}(\mathbf{ki}+\mathbf{ik}) + \epsilon_{xy}(\mathbf{ij}+\mathbf{ji}), \tag{6.02}$$

$$\mathbf{\Psi} = \sigma_{xx}\mathbf{ii} + \sigma_{yy}\mathbf{jj} + \sigma_{zz}\mathbf{kk} + \sigma_{yz}(\mathbf{jk}+\mathbf{kj}) + \sigma_{zx}(\mathbf{ki}+\mathbf{ik}) + \sigma_{xy}(\mathbf{ij}+\mathbf{ji}). \tag{6.03}$$

We suppose the wave normal to be parallel to the $z$ axis. Then the Cartesian equation of the central section of the index ellipsoid perpendicular to the wave normal is obtained by setting $\mathbf{r} = \mathbf{i}x + \mathbf{j}y$ in (2.01) and expanding, with the result

$$n_{xx}^{-2}x^2 + 2n_{xy}^{-2}xy + n_{yy}^{-2}y^2 = 1, \tag{6.04}$$

or in polar coordinates $r^2 = x^2 + y^2$, $\vartheta = \tan^{-1}(y/x)$,

$$n_{xx}^{-2}\cos^2\vartheta + 2n_{xy}^{-2}\sin\vartheta\cos\vartheta + n_{yy}^{-2}\sin^2\vartheta = r^{-2}. \tag{6.05}$$

Setting the derivative of (6.05) with respect to $\vartheta$ equal to zero, we find the angles, $\vartheta = \varphi$, $\varphi + \pi/2$, which the stationary positions (polarizing axes) of the radius vector make with the $x$ axis:

$$\tan 2\varphi = 2n_{xy}^{-2}/(n_{xx}^{-2} - n_{yy}^{-2}). \tag{6.06}$$

Substituting $\vartheta = \varphi$, $\varphi + \pi/2$ in (6.05) we obtain the magnitudes of the stationary values of the radius vector, i.e., the indices of refraction $n_1$, $n_2$ (where $n_1 > n_2$) in the plane of the wave front:

$$n_2^{-2} = n_{xx}^{-2}\cos^2\varphi + n_{yy}^{-2}\sin^2\varphi + n_{xy}^{-2}\sin 2\varphi, \tag{6.07}$$

$$n_1^{-2} = n_{xx}^{-2}\sin^2\varphi + n_{yy}^{-2}\cos^2\varphi - n_{xy}^{-2}\sin 2\varphi. \tag{6.08}$$

Similarly, for the strain dyadic

$$\epsilon_1 = \epsilon_{xx}\cos^2\varphi' + \epsilon_{yy}\sin^2\varphi' + \epsilon_{xy}\sin 2\varphi', \tag{6.09}$$

$$\epsilon_2 = \epsilon_{xx}\sin^2\varphi' + \epsilon_{yy}\cos^2\varphi' - \epsilon_{xy}\sin 2\varphi', \tag{6.10}$$

$$\tan 2\varphi' = 2\epsilon_{xy}/(\epsilon_{xx} - \epsilon_{yy}), \tag{6.11}$$

and, for the stress dyadic,

$$\sigma_1 = \sigma_{xx}\cos^2\varphi'' + \sigma_{yy}\sin^2\varphi'' + \sigma_{xy}\sin 2\varphi'', \tag{6.12}$$

$$\sigma_2 = \sigma_{xx}\sin^2\varphi'' + \sigma_{yy}\cos^2\varphi'' - \sigma_{xy}\sin 2\varphi'', \tag{6.13}$$

$$\tan 2\varphi'' = 2\sigma_{xy}/(\sigma_{xx} - \sigma_{yy}), \tag{6.14}$$

where $\epsilon_1$, $\epsilon_2$ and $\sigma_1$, $\sigma_2$ are secondary principal strains and stresses ($\epsilon_1 > \epsilon_2$, $\sigma_1 > \sigma_2$) in the plane of the wave front, and $\varphi'$ and $\varphi''$ are the angles which the directions of $\epsilon_1$ and $\sigma_1$, respectively, make with the $x$ axis. In an isotropic elastic medium $\varphi = \varphi' = \varphi''$, but this is not necessarily true in a viscoelastic medium.

Returning to the stress-strain-optical equation (4.05) and equating, in turn, the coefficients of the dyads **ii**, **jj**, **ij**, we obtain:

$$n_{xx}^{-2} - n_0^{-2} = \mathbf{R}[\sigma_{xx} - \Theta] + 2\mathbf{S}[\epsilon_{xx} - \theta] + d_1\theta, \tag{6.15}$$

$$n_{yy}^{-2} - n_0^{-2} = \mathbf{R}[\sigma_{yy} - \Theta] + 2\mathbf{S}[\epsilon_{yy} - \theta] + d_1\theta, \tag{6.16}$$

$$n_{xy}^{-2} = \mathbf{R}[\sigma_{xy}] + 2\mathbf{S}[\epsilon_{xy}]. \tag{6.17}$$

Subtracting (6.16) from (6.15),

$$n_{xx}^{-2} - n_{yy}^{-2} = \mathbf{R}[\sigma_{xx} - \sigma_{yy}] + 2\mathbf{S}[\epsilon_{xx} - \epsilon_{yy}]. \tag{6.18}$$

[5] See reference 3, p. 256.

Now, from (6.07), (6.08), (6.17), and (6.18),

$$n_2^{-2}-n_1^{-2}=(n_{xx}^{-2}-n_{yy}^{-2})\cos 2\varphi+2n_{xy}^{-2}\sin 2\varphi,$$
$$=\{\mathbf{R}[\sigma_{xx}-\sigma_{yy}]+2\mathbf{S}[\epsilon_{xx}-\epsilon_{yy}]\}\cos 2\varphi$$
$$+2\{\mathbf{R}[\sigma_{xy}]+2\mathbf{S}[\epsilon_{xy}]\}\sin 2\varphi. \quad (6.19)$$

From (6.09) to (6.14),

$$\epsilon_{xx}-\epsilon_{yy}=(\epsilon_1-\epsilon_2)\cos 2\varphi', \quad (6.20)$$
$$2\epsilon_{xy}=(\epsilon_1-\epsilon_2)\sin 2\varphi', \quad (6.21)$$
$$\sigma_{xx}-\sigma_{yy}=(\sigma_1-\sigma_2)\cos 2\varphi'', \quad (6.22)$$
$$2\sigma_{xy}=(\sigma_1-\sigma_2)\sin 2\varphi'', \quad (6.23)$$

and, hence,

$$n_2^{-2}-n_1^{-2}=\{(\mathbf{R}[(\sigma_1-\sigma_2)\cos 2\varphi'']$$
$$+2\mathbf{S}[(\epsilon_1-\epsilon_2)\cos 2\varphi']\}\cos 2\varphi$$
$$+\{\mathbf{R}[(\sigma_1-\sigma_2)\sin 2\varphi'']$$
$$+2\mathbf{S}[(\epsilon_1-\epsilon_2)\sin 2\varphi']\}\sin 2\varphi. \quad (6.24)$$

*Equation (6.24) is the stress-strain-optical relation in the plane of the wave front.*

In the case of the medium exhibiting instantaneous and retarded elasticity, but no flow ($\eta_3=\infty$, $G_1$ and $G_2$=const.), the partial differential coefficients are absent from the operators **R** and **S**, and (6.24) reduces to

$$2G_1(n_2^{-2}-n_1^{-2})=\left(c_1-c_2-\frac{c_2}{G_2}\frac{d\eta_2}{dt}\right)$$
$$\times(\sigma_1-\sigma_2)\cos 2(\varphi-\varphi'')$$
$$+2G_1c_2\left(1+\frac{1}{G_2}\frac{d\eta_2}{dt}\right)(\epsilon_1-\epsilon_2)\cos 2(\varphi-\varphi'). \quad (6.25)$$

The corresponding equation for a medium exhibiting instantaneous elasticity and flow, but no retarded elasticity, is obtained by setting $c_2=0$ in (6.24), with the result

$$2G_1(n_2^{-2}-n_1^{-2})=c_1(\sigma_1-\sigma_2)\cos 2(\varphi-\varphi''). \quad (6.26)$$

Finally, for an elastic medium, $\varphi=\varphi''$ and (6.26) reduces to

$$2G_1(n_2^{-2}-n_1^{-2})=c_1(\sigma_1-\sigma_2). \quad (6.27)$$

When account is taken of the fact that

$$n_2^{-2}-n_1^{-2}\approx 2(n_1-n_2)/n_0^3, \quad (6.28)$$

Eq. (6.27), and the relation $\varphi=\varphi''$ comprise the usual stress-optical laws for isotropic, elastic media, in which case the determination of $n_1-n_2$, and $\varphi$ gives directly, $\sigma_1-\sigma_2$ and $\varphi''$. For a viscoelastic medium, however, a knowledge of $n_1-n_2$ and $\varphi$ is not sufficient, in general, to determine $\sigma_1-\sigma_2$ and $\varphi''$, as may be seen by inspecting (6.24) or (6.25) or (6.26).

## VII. SIMPLIFICATION WHEN THE STRESS DEVIATION IS A PRODUCT FUNCTION

The complexity of the stress-strain-optical equations is due to the fact that, in general, the secondary principal axes of stress and strain, in the plane of the wave front, do not coincide with each other or with the polarizing axes and also the relative orientations of these axes may vary with time.

Consider a state of stress in which the stress deviation is the product of two functions, one of which depends only on $x$, $y$, and $z$ and the other only on $t$; thus

$$\Psi_d=\bar{\Psi}_d(x, y, z)f(t). \quad (7.01)$$

(In what follows, a bar over a symbol indicates that the quantity is a function of $x, y, z$ only.) As a result of the stress-strain relation (3.16), $\Phi_d$ will then be a product function also; thus

$$\Phi_d=\bar{\Phi}_d(x, y, z)g(t), \quad (7.02)$$

and, accordingly,

$$\bar{\Psi}_d/2\bar{\Phi}_d=\mathbf{Q}[g]/\mathbf{P}[f]=G_0, \quad (7.03)$$

where $G_0$ is an arbitrary constant of the nature and dimensions of a shear modulus. Hence, the assumption (7.01) implies[6]

$$\bar{\Psi}_d=2G_0\bar{\Phi}_d, \quad (7.04)$$
$$\mathbf{Q}[g]=G_0\mathbf{P}[f]. \quad (7.05)$$

Now, from (7.01), (7.02), (6.02), and (6.03),

$$\sigma_{xy}=\overline{\sigma_{xy}}(x, y, z)f(t),$$
$$\sigma_{xx}-\sigma_{yy}=\overline{\sigma_{xx}-\sigma_{yy}}(x, y, z)f(t), \quad (7.06)$$
$$\epsilon_{xy}=\overline{\epsilon_{xy}}(x, y, z)g(t),$$
$$\epsilon_{xx}-\epsilon_{yy}=\overline{\epsilon_{xx}-\epsilon_{yy}}(x, y, z)g(t), \quad (7.07)$$

whence, from (6.11) and (6.14), $\varphi'$ and $\varphi''$ are functions of $x, y, z$ only. Also, from (7.04),

$$\overline{\sigma_{xy}}=2G_0\overline{\epsilon_{xy}}, \quad \overline{\sigma_{xx}-\sigma_{yy}}=2G_0\overline{\epsilon_{xx}-\epsilon_{yy}}, \quad (7.08)$$

so that $\varphi'=\varphi''$. Finally, dividing (6.17) by (6.18)

[6] See reference 1, p. 563, for comparisons.

and using (7.06), (7.07), (7.08), and (6.06), we find

$$\varphi = \varphi' = \varphi'' = \varphi(x, y, z). \tag{7.09}$$

*Thus the condition (7.01) makes the secondary principal axes of birefringence, strain and stress coincide.*

Noting that

$$\sigma_1 - \sigma_2 = \overline{\sigma_1 - \sigma_2} f(t) \quad \text{and} \quad \epsilon_1 - \epsilon_2 = \overline{\epsilon_1 - \epsilon_2} g(t),$$

(6.24) becomes

$$n_2^{-2} - n_1^{-2} = \overline{\sigma_1 - \sigma_2}\mathbf{R}[f] + 2\overline{\epsilon_1 - \epsilon_2}\mathbf{S}[g]. \tag{7.10}$$

Finally, since $\overline{\sigma_1 - \sigma_2} = 2G_0\overline{\epsilon_1 - \epsilon_2}$,

$$n_2^{-2} - n_1^{-2} = \overline{\sigma_1 - \sigma_2}\{\mathbf{R}[f] + \mathbf{S}[g]/G_0\}, \tag{7.11}$$

or, assuming that (6.28) is an equality,

$$n_1 - n_2 = \overline{\sigma_1 - \sigma_2} C(t), \tag{7.12}$$

where

$$C(t) = n_0^3\{\mathbf{R}[f] + \mathbf{S}[g]/G_0\}/2. \tag{7.13}$$

Thus, the condition $\mathbf{\Psi}_d = \overline{\mathbf{\Psi}}_d f(t)$ leads to a simple stress-optical relation for viscoelastic media entirely analogous to that for an elastic medium. The time- and temperature-dependent relative stress-optical coefficient $C(t)$ can be determined in a polariscope by an experiment in simple tension in which the tensile stress $\sigma$ is applied according to the desired law $\bar{\sigma} f(t)$ and the temperature is varied with time in the desired manner.

Similarly a strain-optical relation

$$n_1 - n_2 = \overline{\epsilon_1 - \epsilon_2} c(t) \tag{7.14}$$

may be written, where

$$c(t) = n_0^3\{G_0\mathbf{R}[f] + \mathbf{S}[g]\}. \tag{7.15}$$

The time- and temperature-dependent, relative strain-optical coefficient $c(t)$ may be obtained in a polariscope by an elongation experiment in which the stretch $\epsilon$ follows a prescribed law $\bar{\epsilon} g(t)$ and the temperature is varied with time in the desired manner.

These experiments may be used, also, to solve Eq. (7.05).

## VIII. SOME THEOREMS FROM ELASTICITY THEORY

In the remainder of this paper, photoelasticity will be considered as an experimental method for finding the state of stress in an elastic body, i.e., for solving the equations of elasticity, and we shall study the conditions under which viscoelastic materials may be used as models for this purpose. We have already seen that the condition $\mathbf{\Psi}_d = \overline{\mathbf{\Psi}}_d f(t)$ is sufficient to assure the simple relations (7.09) and (7.12) between birefringence and stress. We require, then,

*A.* The necessary and sufficient conditions on the viscoelastic model for the stress in it to be of this form.

In addition, we require,

*B.* The necessary and sufficient conditions for the stress in the viscoelastic model to correspond to the stress in an elastic body of the same shape under analogous boundary conditions.

The problem of finding these conditions is simplified by the fact that, from the viewpoint of mechanics of continua, elastic and viscoelastic media differ only in the relation between stress and strain. Hence, all of the theorems of elasticity theory which are independent of the stress-strain relation apply as well to viscoelastic bodies. The theorems which we shall need are:[7]

(1) In a body in equilibrium the stress dyadic, in the absence of body forces, satisfies the stress equation of equilibrium:

$$\nabla \cdot \mathbf{\Psi} = 0. \tag{8.01}$$

(In a viscoelastic body, a state of "equilibrium" is one in which the inertia term, in the equation of motion, is negligible.)

(2) The pure small-strain dyadic satisfies the compatibility equation:

$$\nabla \times \Phi \times \nabla = 0. \tag{8.02}$$

(3) In a simply connected body, continuity and single valuedness of strain imply single valuedness of displacement.

(4) In a simply connected body in which the strain is continuous and single valued, the compatibility condition ((8.02)) is a sufficient condition for single valuedness of displacement.

(5) In an $(s+1)$-connected body, in which the strain is continuous and single valued, the displacement may be many valued to the extent of possessing $s$ definite cyclic functions, which may be interpreted as $s$ discontinuities across $s$ sur-

[7] A. E. H. Love, *Theory of Elasticity* (Cambridge University Press, London, 1927), fourth edition, pp. 85, 221–225.

faces $\sigma_i$. The surfaces $\sigma_i$ are those which reduce the multiply-connected body to a simply connected one, and the discontinuities must be of the form

$$\delta\mathbf{u}_i=\mathbf{a}_i+\mathbf{r}\cdot(\mathbf{I}\times\mathbf{b}_i), \quad i=1, 2, \cdots s, \quad (8.03)$$

where $\mathbf{a}_i$ and $\mathbf{b}_i$ are independent of $x, y, z$ and $\mathbf{r}=\mathbf{i}x+\mathbf{j}y+\mathbf{k}z$.

In addition, a uniqueness theorem will be established along lines paralleling closely the Kirchhoff uniqueness theorem of elasticity theory.[8]

## IX. COMPATIBILITY CONDITION

The stress-strain relation for an elastic medium with shear modulus $G$ and Poisson's ratio $\nu$ may be written as

$$2G\mathbf{\Phi}=\mathbf{\Psi}-\alpha\Theta\mathbf{I}, \quad (9.01)$$

where

$$\alpha=3\nu/(1+\nu),$$

and $\nu$ is Poisson's ratio. Hence the compatibility condition (8.02) requires that the stress in an elastic medium satisfy the equation

$$\nabla\times(\mathbf{\Psi}-\alpha\Theta\mathbf{I})\times\nabla=0. \quad (9.02)$$

The stress-strain relations (3.08) and (3.16) for the idealized viscoelastic medium may be written in the form

$$2\mathbf{Q}[\mathbf{\Phi}]=\mathbf{P}[\mathbf{\Psi}-\alpha_1\Theta\mathbf{I}] -(1-\alpha_1)(\mathbf{P}-\mathbf{Q}/G_1)[\Theta\mathbf{I}], \quad (9.03)$$

where

$$\alpha_1=3\nu_1/(1+\nu_1).$$

In any physical experiment the functions $\mathbf{\Phi}$, $\mathbf{\Psi}$, and $\Theta$ will be such that the order of differentiation with respect to $t$ and $x, y, z$ may be reversed.[9] Hence, the compatibility condition requires that the stress in the viscoelastic medium satisfy the equation

$$\mathbf{P}[\nabla\times(\mathbf{\Psi}-\alpha_1\Theta\mathbf{I})\times\nabla] -(1-\alpha_1)(\mathbf{P}-\mathbf{Q}/G_1)[\nabla\times\Theta\mathbf{I}\times\nabla]=0. \quad (9.04)$$

Conditions $B$, Section VIII, require the stress in the viscoelastic medium to satisfy differential equations of the same form as those for an elastic medium. Comparing (9.04) with (9.02), we see that this is true only if $\alpha_1=1$ or if $\Theta$ is independent of $x$, $y$, and $z$. The form of $\Theta$ is usually not known in advance, is difficult to measure, and, moreover, is not likely to be found to be independent of the space coordinates even if measured. Hence, we shall have to require $\alpha_1=1$, i.e., the viscoelastic medium must be incompressible. Fortunately, many viscoelastic materials suitable for photoelasticity have Poisson's ratios nearly equal to one-half.

It should be noted that the stress in an incompressible viscoelastic medium satisfies a differential equation ((9.04) with $\alpha_1=1$) of the same form as the differential equation ((9.02) with $\alpha=1$) governing the stress in an *incompressible* elastic medium.[10]

## X. NECESSARY AND SUFFICIENT CONDITIONS FOR $\mathbf{\Psi}_d=\overline{\mathbf{\Psi}}_d f(t)$ IN A SIMPLY CONNECTED BODY

We note, first of all, that in an incompressible medium $\mathbf{\Phi}=\mathbf{\Phi}_d$ and, since $\mathbf{I}:\mathbf{\Phi}_d=0$,

$$\mathbf{\Psi}:\mathbf{\Phi}=(\mathbf{\Psi}_d+\Theta\mathbf{I}):\mathbf{\Phi}_d=\mathbf{\Psi}_d:\mathbf{\Phi}_d. \quad (10.01)$$

Also, the pure strain dyadic $\mathbf{\Phi}$ is the self-conjugate part of the strain dyadic $\nabla\mathbf{u}$, where $\mathbf{u}$ is the displacement vector. Thus

$$2\mathbf{\Phi}=2\mathbf{\Phi}_d=\nabla\mathbf{u}+\mathbf{u}\nabla. \quad (10.02)$$

Then, since $\mathbf{\Psi}$ is also self-conjugate,

$$\mathbf{\Psi}:\mathbf{\Phi}=\tfrac{1}{2}\mathbf{\Psi}:(\nabla\mathbf{u}+\mathbf{u}\nabla)=\mathbf{\Psi}:\nabla\mathbf{u} =\nabla\cdot(\mathbf{\Psi}\cdot\mathbf{u})-(\nabla\cdot\mathbf{\Psi})\cdot\mathbf{u}. \quad (10.03)$$

But, in a body in equilibrium, $\nabla\cdot\mathbf{\Psi}=0$. Hence, in an incompressible body in equilibrium,

$$\mathbf{\Psi}_d:\mathbf{\Phi}_d=\nabla\cdot(\mathbf{\Psi}\cdot\mathbf{u}), \quad (10.04)$$

and we may write

$$\int\int\int\mathbf{\Psi}_d:\mathbf{\Phi}_d d\tau=\int\int\int\nabla\cdot(\mathbf{\Psi}\cdot\mathbf{u})d\tau, \quad (10.05)$$

where $d\tau$ is an element of volume and the integrations are extended throughout the volume of the viscoelastic body under consideration.

As noted in Section VIII, in a *simply connected* body no displacement discontinuities are possible if the strain is continuous and single valued.

[8] See reference 7, p. 170.

[9] It is assumed that the temperature distribution throughout the body is homogeneous, so that $\eta_i$ and $G_i$ depend only on $t$ and not on $x$, $y$, or $z$.

[10] See reference 1, p. 561

In such a body the divergence theorem yields

$$\int\int\int \nabla\cdot(\Psi\cdot\mathbf{u})d\tau=\int\int \mathbf{n}\cdot\Psi\cdot\mathbf{u}d\sigma, \quad (10.06)$$

where $d\sigma$ is an element of the surface, $\mathbf{n}$ is its unit, outward-drawn, normal vector and the surface integral is extended over the whole surface of the body. (In the surface integral, $\mathbf{u}$ is the displacement at the surface of the body and the traction

$$\mathbf{T}=\mathbf{n}\cdot\Psi \quad (10.07)$$

is the surface traction.) Hence,

$$\int\int\int \Psi_d:\Phi_d d\tau=\int\int \mathbf{T}\cdot\mathbf{u}d\sigma. \quad (10.08)$$

We have seen (Section VII), that, if

$$\Psi_d=\overline{\Psi}_d f(t),$$

$\Phi_d$ becomes $\bar{\Phi}_d g(t)$. Hence, from (10.08), a necessary condition for $\Psi_d$ to be of the form $\overline{\Psi}_d f(t)$ is

$$\mathbf{T}\cdot\mathbf{u}=\bar{\mathbf{T}}\cdot\bar{\mathbf{u}}f(t)g(t), \quad (10.09)$$

and further, in view of (10.02),

$$\mathbf{T}=\bar{\mathbf{T}}(x, y, z)f(t), \quad (10.10)$$

$$\mathbf{u}=\bar{\mathbf{u}}(x, y, z)g(t) \quad (10.11)$$

on the boundary. This excludes, for example, *moving loads*.

Equation (10.09) is not a sufficient condition for $\Psi_d$ to be of the required form; for there may be present another stress system $\Psi^0$ (an "initial stress") for which $\mathbf{n}\cdot\Psi^0\cdot\mathbf{u}^0=0$ on the boundary. The conditions are completed by requiring $\Psi_d{}^0$ to be either zero or of the form

$$\Psi_d{}^0=\overline{\Psi}_d{}^0(x, y, z)f(t). \quad (10.12)$$

The presence of initial stress is easily detected optically but is usually difficult to measure. Generally it will not be of the form (10.12) unless it is a *residual* stress remaining from a previous experiment with boundary conditions (10.09), in which case the calibration test (Section XIII) must include the complete history of $f(t)$, $g(t)$, $\eta_i(t)$, and $G_i(t)$ from the unstressed state.

## XI. UNIQUENESS THEOREM FOR A SIMPLY CONNECTED BODY

From (10.08), (10.10), (10.11), (7.01), (7.02), and (7.04),

$$\int\int\int \overline{\Psi}_d:\overline{\Psi}_d d\tau=2G_0\int\int \bar{\mathbf{T}}\cdot\bar{\mathbf{u}}d\sigma. \quad (11.01)$$

Let $\Psi_d{}^I$ and $\Psi_d{}^{II}$ be the stress deviations, vanishing at $t=0$, of two equilibrium states of stress $\Psi^I$ and $\Psi^{II}$ in an incompressible body. We suppose the tractions $\mathbf{T}^I$, $\mathbf{T}^{II}$ and the displacements $\mathbf{u}^I$, $\mathbf{u}^{II}$ to be of such form that, on the boundary,

$$\mathbf{T}^I\cdot\mathbf{u}^I=\bar{\mathbf{T}}^I\cdot\bar{\mathbf{u}}^I f(t)g(t)$$

and

$$\mathbf{T}^{II}\cdot\mathbf{u}^{II}=\bar{\mathbf{T}}^{II}\cdot\bar{\mathbf{u}}^{II}f(t)g(t).$$

Then, by Section X,

$$\Psi_d{}^I=\overline{\Psi}_d{}^I f(t), \quad \Psi_d{}^{II}=\overline{\Psi}_d{}^{II}f(t).$$

Let

$$\Psi^*=\Psi^I-\Psi^{II},$$

$$\Psi_d{}^*=\overline{\Psi}_d{}^I-\overline{\Psi}_d{}^{II},$$

$$\bar{\mathbf{T}}^*\cdot\bar{\mathbf{u}}^*=(\bar{\mathbf{T}}^I-\bar{\mathbf{T}}^{II})\cdot(\bar{\mathbf{u}}^I-\bar{\mathbf{u}}^{II}).$$

Then, by (11.01),

$$\int\int\int \overline{\Psi}_d{}^*:\overline{\Psi}_d{}^* d\tau=2G_0\int\int \bar{\mathbf{T}}^*\cdot\bar{\mathbf{u}}^* d\sigma. \quad (11.02)$$

Now, if

$$\bar{\mathbf{T}}^*\cdot\bar{\mathbf{u}}^*=0 \quad (11.03)$$

on the boundary, the surface integral in (11.02) vanishes and hence the volume integral also vanishes. But $\overline{\Psi}_d{}^*:\overline{\Psi}_d{}^*$ is equal to the sum of the squares of the differences between the principal stresses of $\Psi^*$ and its mean tension. Hence, $\overline{\Psi}_d{}^*:\overline{\Psi}_d{}^*$ is always positive or zero. Thus, if the volume integral in (11.02) vanishes, $\Psi_d{}^*:\Psi_d{}^*=0$ and, consequently, $\Psi^I$ and $\Psi^{II}$ can differ only by a mean tension.

A mean tension produces no relative birefringence and cannot, in general, be determined from measurements of relative phase retardation and orientation of polarizing axes. An exception is when one or more of the principal stresses is known, *a priori*, to vanish, as at a point on a free boundary or in a state of plane stress. Then, however, the condition $\overline{\Psi}_d{}^*:\overline{\Psi}_d{}^*=0$ makes $\Psi^I=\Psi^{II}$.

Let

$$\mathbf{T}=T_\lambda\boldsymbol{\lambda}+T_\mu\boldsymbol{\mu}+T_\nu\boldsymbol{\nu}, \quad \mathbf{u}=u_\lambda\boldsymbol{\lambda}+u_\mu\boldsymbol{\mu}+u_\nu\boldsymbol{\nu},$$

where $\boldsymbol{\lambda}$, $\boldsymbol{\mu}$, $\boldsymbol{\nu}$ constitute a set of mutually perpendicular unit vectors at each point of the boundary. Then (11.03) may be written as

$$(\bar{T}_\lambda{}^I-\bar{T}_\lambda{}^{II})(\bar{u}_\lambda{}^I-\bar{u}_\lambda{}^{II})+(\bar{T}_\mu{}^I-\bar{T}_\mu{}^{II})(\bar{u}_\mu{}^I-\bar{u}_\mu{}^{II})+(\bar{T}_\nu{}^I-\bar{T}_\nu{}^{II})(\bar{u}_\nu{}^I-\bar{u}_\nu{}^{II})=0, \quad (11.04)$$

where the superscripts and bars have the same significance as before. Hence, the specification of any one of eight combinations of boundary conditions (one traction, one displacement and six mixed conditions) will make the two states of stress identical except, possibly, for a mean tension.

## XII. EXTENSION TO MULTIPLY CONNECTED BODIES

As stated in Section VIII, in an $(s+1)$-connected body there may be $s$ displacement discontinuities $\delta\mathbf{u}_i$, of the form (8.03), across $s$ surfaces $\sigma_i$ which reduce the multiply connected body to a simply connected one. In this case the divergence theorem, instead of having the form (10.06), becomes

$$\int\int\int\nabla\cdot(\boldsymbol{\Psi}\cdot\mathbf{u})d\tau=\int\int\mathbf{n}\cdot\boldsymbol{\Psi}\cdot\mathbf{u}d\sigma+\sum_{i=1}^{s}\int\int\mathbf{n}_i\cdot\boldsymbol{\Psi}\cdot\delta\mathbf{u}_i d\sigma_i, \quad (12.01)$$

where the surface integrals under the summation sign are taken over the surfaces of discontinuity.

Since the additional terms in (12.01) are of the same form as the integral over the external surface, all that has been said above for the quantity $\mathbf{n}\cdot\boldsymbol{\Psi}\cdot\mathbf{u}$ applies also to each of the quantities $\mathbf{n}_i\cdot\boldsymbol{\Psi}\cdot\delta\mathbf{u}_i$.

The additional stress introduced by the "dislocations" $\delta\mathbf{u}_i$ may be looked upon from two points of view. First, the additional stress may be regarded as an initial stress and treated in the same manner as $\boldsymbol{\Psi}^0$ above. Second, it may be required to introduce dislocations deliberately, as in experiments to determine the effect of a difference in Poisson's ratio between model and prototype or in applications of Biot's analogy between thermal stress and dislocations. In such cases the only additional condition that must be satisfied is that the $\delta\mathbf{u}_i$ be of the form $\delta\bar{\mathbf{u}}_i g(t)$.

## XIII. RECAPITULATION AND REMARKS

For an $(s+1)$-connected, viscoelastic body in a state of equilibrium, the necessary and sufficient conditions, for the stress deviation to be of the form $\overline{\boldsymbol{\Psi}}_d f(t)$ and for $\overline{\boldsymbol{\Psi}}_d$ to correspond to the stress in an elastic body of the same shape under analogous boundary conditions, are

(1) The viscoelastic material must be incompressible.**

(2) As in elasticity theory, over each and every portion of the surface of the body and the $s$ surfaces of discontinuity of displacement, one of the following eight combinations of boundary conditions must be prescribed:

$T_\lambda, T_\mu, T_\nu$; $T_\lambda, T_\mu, u_\nu$; $T_\lambda, u_\mu, T_\nu$; $u_\lambda, T_\mu, T_\nu$; $u_\lambda, u_\mu, u_\nu$; $u_\lambda, u_\mu, T_\nu$; $u_\lambda, T_\mu, u_\nu$; $T_\lambda, u_\mu, u_\nu$.

(Note that different "mixtures" may be prescribed over different portions of the surfaces. The first three mixed conditions consist of a component of displacement and the resultant traction in the plane perpendicular to it, while the second three mixed conditions consist of a component of traction and the resultant displacement in the plane perpendicular to it.)

(3) Components of boundary traction, where prescribed, must be of the form

$$\bar{T}_j(x, y, z)f(t), \quad j=\lambda, \mu, \nu$$

and components of boundary displacement, where prescribed, must be of the form

$$\bar{u}_j(x, y, z)g(t), \quad j=\lambda, \mu, \nu,$$

where $\mathbf{Q}[g]=G_0\mathbf{P}[f]$.

(4) The initial stress must be either zero or of the form $\overline{\boldsymbol{\Psi}}_d{}^0 f(t)$, in which case $f(t)$ must be reckoned from a zero state of stress.

Under these conditions the stress and strain in the viscoelastic body are related to the birefringence through the simple formulas

$$\varphi=\varphi'=\varphi'',$$

$$n_1-n_2=\overline{\sigma_1-\sigma_2}C(t),$$

$$n_1-n_2=\overline{\epsilon_1-\epsilon_2}c(t),$$

and the barred components of stress and strain

** In practice, Poisson's ratio should differ little from $\frac{1}{2}$.

are those for an incompressible elastic body with shear modulus $G_0$.

The time- and temperature-dependent stress- and strain-optical coefficients $C$ and $c$ may be determined by experiments on a simple-tension specimen of the same material as the model. For example, the viscoelastic model, held at constant temperature, may be loaded by several forces whose directions and points of application on the boundary are fixed, but whose magnitudes, while maintaining constant ratios, vary with time according to a law $f(t)$. A photographic record of the fringe pattern is made at a time $t_1$. This is the most common type of photoelastic test. (The time-dependence may be introduced simply by the process of reaching and maintaining constant loads.) The calibration bar must be loaded according to the same time-law, at the same temperature, and the ratio $(n_1-n_2)/\overline{\sigma_1-\sigma_2}$ calculated at time $t_1$. This gives the stress-optical coefficient $C(t_1)$ for the photograph.

It is interesting to notice that if the temperature is varied during the model test, as in the freezing method of three-dimensional photoelasticity, the only *additional* requirement is that, when finding the calibration factor $C(t_1)$, the calibration specimen must be subjected to the same temperature-time cycle as the model. Thus, the freezing method is not restricted to models made of cross-linked polymers.

In the case of mixed, non-zero, boundary conditions a preliminary experiment must be performed, for example, in simple tension, to determine the $g(t)$ corresponding to a given $f(t)$ or vice versa.

In the determination of stress-concentration factors it is not necessary to know the stress-optical coefficient when using elastic materials or viscoelastic materials in which $\eta_2$ and $\eta_3$ are so large that "optical creep" is negligible. Similarly, it is not necessary to find $C(t_1)$ in stress-concentration factor tests with viscoelastic materials that *do* exhibit optical creep because, if the above conditions are met, the ratio of the stress at any two points of the model is a constant, equal to the corresponding ratio for the analogous elastic body. This is true regardless of flow or of temperature variation with time.

In the transfer, from model to prototype, of results obtained with a model having a Poisson's ratio different fron that of the prototype, the same difficulties are encountered with viscoelastic models as with elastic models. The few theorems governing such transfers that are available for elastic models are applicable to viscoelastic models.

## Calendar of Meetings

**February**

13–17 American Institute of Mining and Metallurgical Engineers, San Francisco, California
18–19 Ninth Pittsburgh Conference on Applied Spectroscopy, Pittsburgh, Pennsylvania
21–24 Technical Association of the Pulp and Paper Industry, New York, New York
26 American Mathematical Society, New York, New York, and Chicago, Illinois
28–March 4 American Society for Testing Materials, Chicago, Illinois

**March**

3–4 Illuminating Engineering Society (South Pacific Regional Conference), San Francisco, California
6–10 American Institute of Chemical Engineers, Los Angeles, California
7–10 Institute of Radio Engineers, New York, New York
8–10 Society of Automotive Engineers, Detroit, Michigan
9 Inter-Society Color Council, New York, New York
10–11 Illuminating Engineering Society (Pacific Northwest Regional Conference) Portland, Oregon
10–12 American Physical Society, Cleveland, Ohio
10–12 Optical Society of America, New York, New York
14–17 American Association of Petroleum Geologists, St. Louis, Missouri
14–17 Chicago Technical Conference and Production Show, Chicago, Illinois
18 Institute of Aeronautical Sciences (Flight Propulsion Meeting), Cleveland, Ohio
27–April 1 American Chemical Society, San Francisco, California
28–30 Society of Automotive Engineers, Cleveland, Ohio

**April**

1–2 American Mathematical Society, Durham, North Carolina
3–5 Illuminating Engineering Society (Southwestern Regional Conference), Houston, Texas
4–8 Society of Motion Picture Engineers, New York, New York
11–12 American Institute of Electrical Engineers Conference on the Industrial Application of Electron Tubes, Buffalo, New York
11–13 Society of Automotive Engineers, New York, New York
11–15 American Society for Metals, Los Angeles, California
16–17 American Physical Society (Southeastern Section) Clemson, South Carolina
19–21 American Institute of Electrical Engineers (Southwest District Meeting), Dallas, Texas
19–22 American Society of Biological Chemists, Detroit, Michigan
20–22 American Geophysical Union, Washington, D. C.
20–23 American Society of Civil Engineers, Oklahoma City, Oklahoma
21–23 American Philosophical Society, Philadelphia, Pennsylvania
24–28 American Ceramic Society, Cincinnati, Ohio
25–27 National Academy of Sciences, Washington, D. C.
25–28 Technical Association of the Pulp and Paper Industry, Grand Rapids, Michigan
28–29 Illuminating Engineering Society (Canadian Regional Conference), Toronto, Ontario
28–30 American Physical Society, Washington, D. C.
29–30 American Mathematical Society, Philadelphia, Pennsylvania and Lawrence, Kansas
30 American Mathematical Society, Stanford University, California

# Air Friction on Rapidly Moving Surfaces*

A. R. KUHLTHAU**
*Rouss Physical Laboratory, University of Virginia, Charlottesville, Virginia*
(Received August 25, 1948)

The friction on a surface moving through air at high speed has been measured over a pressure range from 0.5 to 200 microns. A rotating cylinder type apparatus was used in which the inner cylinder was rotated at very high speed by means of an air turbine arrangement similar to that used in an air driven vacuum type ultracentrifuge. The outer cylinder completely surrounded the inner cylinder and was supported coaxially with it by means of a torsion fiber so that the frictional torque on the rotating cylinder could be measured with precision. Observations were made at rotor speeds varying from 400 r.p.s. to 1700 r.p.s. which corresponds to a linear velocity at the periphery of the rotor varying from 400 to 1700 ft./sec. It was found that at all pressures the observed torque varied in an essentially linear fashion with the rotor speed. The relationship between torque and pressure at constant speed yielded a smooth curve indicative of ordinary viscosity, and this general form held for all speeds tested. However, the results could be explained by kinetic theory treatment of viscosity only at pressures lower than 10 microns; i.e., below the point where the intermolecular collision frequency predominates over that of the collisions between the molecules and the walls.

## INTRODUCTION

THE gaseous friction on moving surfaces has been measured with precision by a number of different workers, especially in the gaseous pressure range from one-half atmosphere up to several atmospheres. However, relatively few measurements have been made of the friction on very rapidly moving surfaces in highly rarefied gases. This paper describes an attempt to measure the friction on a surface moving at high speed in air over the pressure range from 0.5 to 200 microns; i.e., over the range where the mean free paths of the molecules are long in comparison to the dimensions of the apparatus to where they are very short. Recent measurements by Beams and his co-workers[1] of the gaseous friction on rapidly rotating spherical rotors in air at pressures of the order of $10^{-6}$ mm of Hg have demonstrated the feasibility of using high speed rotors for this purpose, and so the high speed rotor technique was chosen for this work.

## APPARATUS

The apparatus used is similar in principle to the rotating cylinder viscosimeter such as has been used by a number of workers for the precise determination of the coefficients of viscosity and slip. In the experiments described in this paper the inner cylinder is rotated by an air driven turbine arrangement similar to that used in the air driven vacuum type ultracentrifuge[2] so that very high rotor speeds could be obtained in gases over the pressure range desired. The outer cylinder was suspended coaxially with the inner rotating cylinder by a phosphor-bronze fiber so that the torque developed could be measured with high precision. Preliminary experiments showed that with the high rotational speeds it is necessary to almost completely surround the rotating cylinder with the outer cylinder in order to avoid errors caused by "leakage of momentum." As a result, ends were placed on the outer cylinder in order to avoid the above mentioned losses. The measurements consisted of determining the torque on the outer cylinder as a function of the air pressure and the speed of the inner rotor.

Figure 1 shows a diagram of the apparatus. The rotor, $R$, machined from a solid block of forged Duralumin alloy ST-14, is four inches in diameter and two inches in length. It is capable of withstanding stresses produced by rotational speeds of over 2000 revolutions per second. The rotor was clamped as shown to a flexible steel shaft 0.100 inch in diameter. The length to diameter ratio of the rotor was made less than

* The work described in this paper has been supported by the Bureau of Ordnance, U. S. Navy under Contract NOrd-7873.
** Now at University of New Hampshire.
[1] Beams, Young, and Moore, J. App. Phys. **17**, 886 (1946).
[2] See for example: J. W. Beams and F. W. Linke, Rev. Sci. Inst. **8**, 160 (1937); J. W. Beams, J. App. Phys. **8**, 795 (1937).

unity in order to obtain a high degree of rotor stability.

The outer cylinder, *B*, was made of brass and was five inches in diameter with a wall thickness of 1/32 inch. It was suspended coaxially with the rotor by means of the phosphor-bronze wire, *F*, 0.0147 inch in diameter. The wire was fastened to the axis of the outer cylinder by clamping it in a close fitting hole with a No. 0-80 set screw. The top end of the suspension wire was clamped in a standard pin vise which was mounted in a tapered plug, *C*. The wire was made to pass through a small (0.0157″) hole at the bottom of the tapered plug so as to insure its being clamped properly in the pin vise. The plug, *C*, was fitted to a tapered hole in a supporting column, which was made to coincide with the axis of rotation of the rotor within 1/1000 inch by means of the accurately machined jacket, *J*, as shown. The jacket consisted of a base plate, cylindrical side and top plate to which the supporting column was fastened. This column was centered in the top plate by means of a shoulder (not shown). Since *C* is centered by this precision construction, the outer cylinder will be coaxial with the rotor provided it is symmetrical about the axis of the fiber, and provided the base plate of the jacket is horizontal. Both of these conditions were fulfilled.

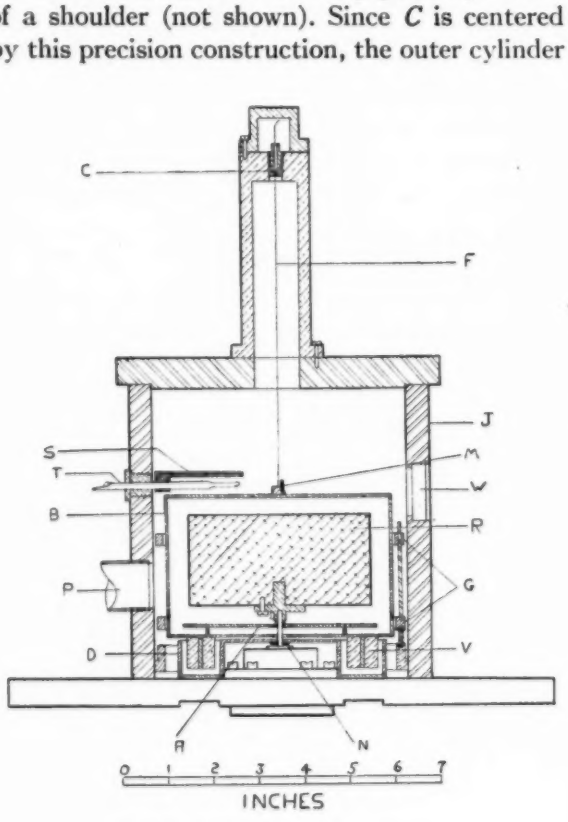



FIG. 1. Section view of apparatus.

The hole in the bottom of the outer cylinder to allow passage of the rotor shaft was 23/64 inch in diameter. The outer cylinder was surrounded by a pair of guard rings, *G*, which restricted the horizontal motion of the cylinder and thus prevented the possibility of contact between the rotor or the rotor shaft and the outer cylinder. The radial clearance between the guard rings and the outer cylinder was 3/32 inch. It was found necessary, however, to damp out the small oscillations of the outer cylinder set up by the necessarily rapid acceleration of the rotor through low speed critical vibrations. This was accomplished by fastening two X-shaped damping vanes, *V*, symmetrically to the bottom of the cylinder. These were allowed to run in an annular dashpot, *D*. Extremely low vapor pressure oil was used as the damping fluid, care being taken that all the residual air and water vapor was removed by evacuating it before it was placed in the dashpot. The dashpot annulus was divided into two sections by means of two radial partitions in order to guard against circulation of the oil around the annulus. The center of the annular dashpot was closed over by a thin brass disk which was fitted to within 1/32 inch of the rotor shaft. To the bottom of this disk was fastened a thin Neoprene washer, *N*, fitting snugly around the rotor shaft. This ensemble thus enclosed the bearing or vacuum gland located in the base plate, and prevented any oil spray from the gland striking the bottom of the outer cylinder. Experiments also demonstrated that it was necessary to equip the outer cylinder with a false bottom or baffle plate, *A*, in order to eliminate the molecular pumping action between the lower end surfaces of the rotor and outer cylinder.

The jacket was equipped with a glass window, *W*, 1½ inches in diameter, so that the deflection of the outer cylinder could be observed by means of an external telescope. Two methods of making observations were employed. The upper periphery of the outer cylinder was marked to read directly in degrees, and so for large deflections the readings were taken directly from this scale. The second method, used in conjunction with smaller

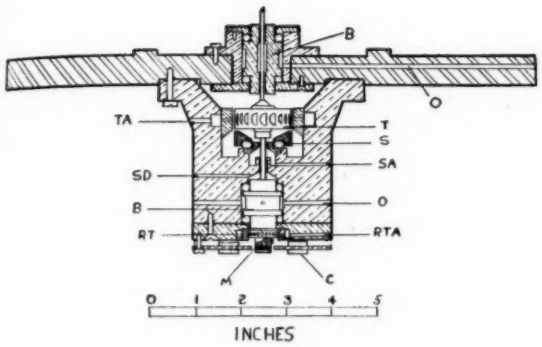



FIG. 2. Drive and support mechanism.

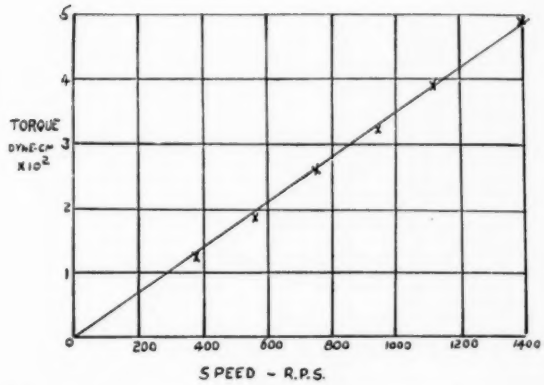



FIG. 3. Torque *vs.* rotor speed. Pressure = 0.48 micron. Temperature = 27°C. ——— Theoretical values. X—X—X Experimental values.

deflections, made use of a small mirror, *M*, mounted at the top of the outer cylinder and as near to the axis as possible. Deflections were then read on an external scale by means of the mirror and the telescope.

The driving and supporting mechanism for the Beams type ultracentrifuge is thoroughly described in the literature.[2,3] Figure 2 shows this driving mechanism in section, and its operation is briefly as follows. The rotor is driven by a Duralumin turbine, *T*, fastened to the rotor shaft by means of a collet arrangement. The rotor shaft is held by two bearings or oil glands, *B*, one above and one below the turbine. The main body of the gland is usually made of brass, a piece of brass rod being turned to the shape shown in the figure. Hard Babbitt is used as the bearing surface, and the proper clearance is such that the polished shaft will just slip through the bearing without forcing. The glands are lubricated with vacuum pump oil (Cenco No. 93050), the oil being forced into the glands through the channels, 0, under about 8 p.s.i. air pressure. The main body of the gland is mounted in round Neoprene rings as shown in the figure in order to provide flexibility. The driving turbine, *T*, was $1\frac{7}{32}$ inches in diameter and $\frac{1}{2}$ inch thick. It is supported on the air cushion formed between its flat under surface and the Bakelite supporting collar, *S*. *S* in turn is flexibly supported on a round Neoprene ring as shown. The supporting air enters through the channel *SA* and is then directed up along the shaft and out over *S*.

In the experiments it was necessary to know the pressure within the rotor chamber, the temperature of the air and the speed of the rotor.

[3] J. W. Beams, Rev. Sci. Inst. 9, 413 (1938).

Temperature was measured by inserting a mercury thermometer, *T*, through the jacket wall (see Fig. 1). The thermometer bulb was located about 3/16 inch above the top of the outer cylinder, and was shielded by a semicircular cover, *S*. *S* was made of Polystyrene, and lined on the inside and outside with aluminum foil so as to make a reflecting cover which enclosed the top of the thermometer over an angle of about 200 degrees. The basic method used to determine the speed of the rotor is as follows. A signal equal in frequency to the speed of the rotor is obtained, amplified and placed on the vertical plates of a cathode-ray oscilloscope. A signal of variable frequency is obtained from a calibrated audio oscillator and put on the horizontal plates of the oscilloscope. The rotor speed is thus obtained directly from the dial of the oscillator by ob-

TABLE I.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| No. of observations | Pressure microns | Approx. mean free path cm | Speed r.p.s. | Mean ex. torque dyne-cm $\times 10^2$ | Ave. deviation | Th. torque dyne-cm $\times 10^2$ |
| 7 | 1 | 7.6 | 375 | 2.52 | 0.05 | 2.71 |
| 8 | 1 | 7.6 | 565 | 3.89 | 0.04 | 4.08 |
| 12 | 1 | 7.6 | 750 | 5.23 | 0.04 | 5.42 |
| 8 | 1 | 7.6 | 935 | 6.51 | 0.03 | 6.76 |
| 11 | 1 | 7.6 | 1120 | 7.85 | 0.03 | 8.10 |
| 8 | 1 | 7.6 | 1310 | 9.16 | 0.03 | 9.47 |
| 3 | 1 | 7.6 | 1400 | 9.83 | 0.03 | 10.1 |
| 3 | 0.48 | 15.2 | 375 | 1.23 | 0.01 | 1.30 |
| 4 | 0.48 | 15.2 | 565 | 1.92 | 0.03 | 1.96 |
| 8 | 0.48 | 15.2 | 750 | 2.60 | 0.03 | 2.61 |
| 9 | 0.48 | 15.2 | 935 | 3.28 | 0.04 | 3.25 |
| 5 | 0.48 | 15.2 | 1120 | 3.97 | 0.04 | 3.88 |
| 4 | 0.48 | 15.2 | 1310 | 4.66 | 0.08 | 4.53 |
| 2 | 0.48 | 15.2 | 1400 | 4.91 | — | 4.85 |

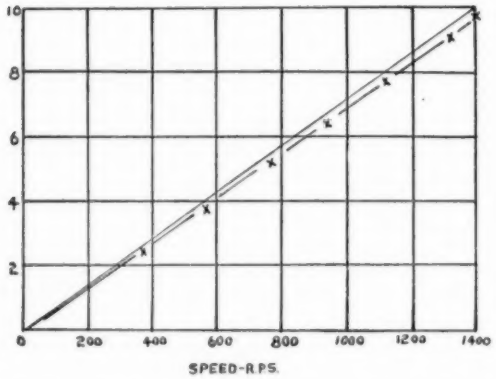



FIG. 4. Torque *vs.* rotor speed. Pressure = 1 micron. Temperature = 26°C. ——— Theoretical values. X—X—X Experimental values.

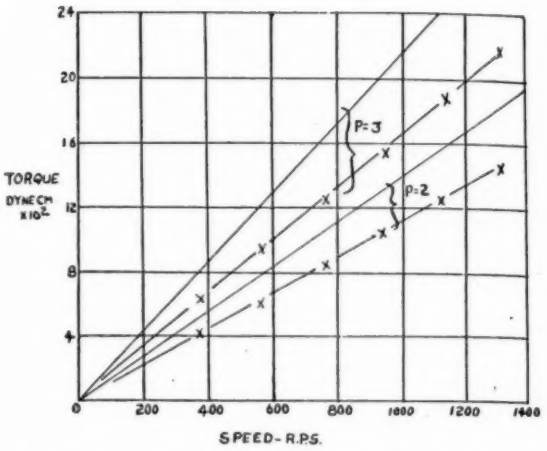



FIG. 5. Torque *vs.* rotor speed. $P$ = pressure in microns. Temperature = 26°C. ——— Theoretical values. X—X—X Experimental values.

serving the resulting Lissajous figure on the oscilloscope screen. The rotor signal was obtained either by mounting a small magnet on the rotor shaft and causing it to rotate within a pair of "pick-up" coils, or by picking up the fundamental note of the rotor by means of a microphone. The former arrangement is shown in Fig. 2, the small magnet, $M$, rotating within the coils, $C$. The pressure within the rotor chamber was controlled by leaking a continuous stream of dry air into the system against the pumps by means of an adjustable capillary leak. A McLeod gauge was used as a standard to measure the existing pressure, and a continuous indication of the pressure was obtained from either a Knudsen absolute manometer, or a Pirani gauge, depending upon the pressure range involved.

## EXPERIMENTAL RESULTS

Experiments were conducted at rotor speeds ranging from about 400 revolutions per second to about 1700 revolutions per second. This corresponds to a linear velocity at the periphery of the rotor ranging from approximately 400 ft./sec. to 1700 ft./sec. Dry air was used as the gaseous medium, and observations were made over a pressure range from below 0.5 micron to over 100 microns. Two separate procedures were used in obtaining the data at the various speeds and pressures. The first was to hold the pressure fixed and vary the speed. Observations were usually made at increasing speed increments of 100 revolutions per second over the entire speed range, and then the readings were repeated at equal descending increments. After this the pressure was changed to a new value and the entire process repeated. The second procedure consisted of holding the speed constant and varying the pressure over a range appropriate to the leak being used. The observations were usually made first with ascending pressure increments, and then with the same descending increments, after which the rotor speed was changed to a new value and the whole process repeated. Identical results were obtained with either procedure and therefore, the former was the one most commonly used as equilibrium was attained more rapidly after speed changes than after pressure changes.

For all pressures below about 1 micron the mean free path in the gas is longer than the smallest restricting dimension of the apparatus. This means that the intermolecular collision frequency is small compared to the frequency of collisions between the molecules and the surfaces of the apparatus, and hence this range has often been appropriately referred to as the "free molecule range." The retarding torque on the rotor which is to be expected in this range may readily be calculated from simple kinetic theory considerations. If one assumes that the gas in the system behaves as an ideal gas with a Maxwellian distribution of velocities, and that all molecules that strike the rotor are diffusely reflected, the expression for the retarding torque will be

$$\text{Torque} = \pi p(M/2\pi RT)^{\frac{1}{2}}\omega a^3(a+2l), \qquad (1)$$

where $p$ = pressure in the system, $\omega$ = angular velocity, $M$ = molecular weight of the gas under consideration, $T$ = the Absolute temperature, $a$ = radius of the rotor, $l$ = length of the rotor, $R$ = the gas constant. For air and the dimensions of the particular apparatus used, this becomes

$$\text{Torque} = 1.49(p\omega/T^{\frac{1}{2}}).$$

This then is the torque applied to the outer cylinder, which is equal to the measured restoring torque of the suspension.

A large number of observations were made in this range, and the results are tabulated in Table I. The theoretical values are calculated from Eq. (1), and the experimental values are the mean of a number of observations made at the particular pressure and speed. Column 5 gives the average deviation from the mean. The plot of these data are shown in Figs. 3 and 4. In each case the temperature was constant within 1 degree centigrade, and so the torque is plotted against speed at constant pressure. The experimental results obtained at a pressure of 0.48 micron are seen to be in excellent agreement with the "free molecule theory," but at a pressure of 1 micron they are beginning to deviate from this theory, although the deviation is not yet more than a few percent. In Fig. 5 similar curves are plotted from data obtained at pressures of 2 and 3 microns, and the deviation between experimental and theoretical values is seen to be even more marked. Several observations were made at a pressure of 0.25 micron, and the measured torque agreed with the theoretical value at all speeds within $\pm 5$ percent which was the limit of experimental accuracy at this pressure. From the above data it seems reasonable to conclude that the "free molecule theory" as expressed by Eq. (1) will account for the retarding torque on the rotor at pressures below 1 micron for all speeds tested, but that this theory no longer holds above 1 micron.

The experiments were continued into the range of pressures from 1 to 100 microns in order to observe the nature of the retarding torque on the rotor in a region where the "free molecule

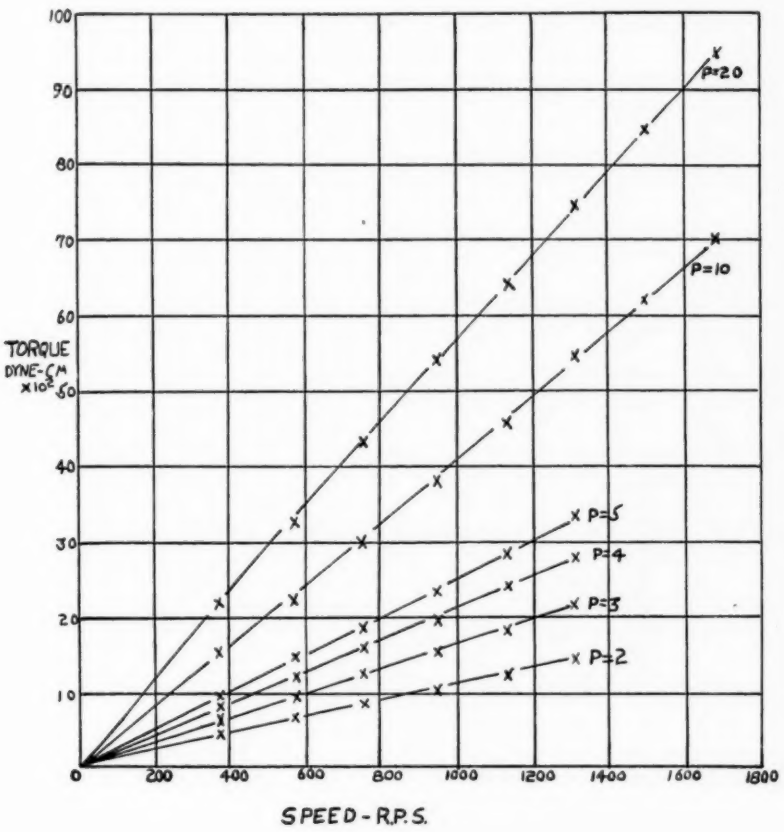



FIG. 6. Torque *vs.* rotor speed. $P$ = pressure in microns. Temperature = 26°C.

viscosity" does not hold. Data were obtained at various pressures within this range and at speeds varying from 375 revolutions per second to 1695 revolutions per second. During the experiments the temperature of the system remained within a degree or two of 26°C at all speeds for pressures up to 30 microns. At higher pressures there was a slight rise in temperature for the higher speeds. A considerable number of observations were thus made at temperatures varying from 26°C to 40°C. However, the data obtained were not sufficient to permit an exact determination of the observed torque as a function of temperature, but all indications were that to a close approximation the torque was directly proportional to the Absolute temperature. Consequently, in assembling the data, all observed values were corrected to the appropriate value at 26°C.

Figures 6 and 7 show a plot of torque against rotor speed using the pressure, $p$, as a parameter. As can be seen the plots are all straight lines within very close limits at pressures up to 20 microns. However, at pressures higher than 20 microns the curves slope upwards slightly at the higher speeds, as can be seen from Fig. 7. An analysis of the values of the ratio of torque *vs.* rotor speed for these curves shows that the slope increases at the higher pressures, the percent increase over the entire range of speeds varying from 3 percent at 30 microns to about 7 percent at 100 microns. These deviations from a straight line are of the same order as the experimental errors at these higher pressures, and hence no conclusions can be drawn concerning the non-linearity of the curves.

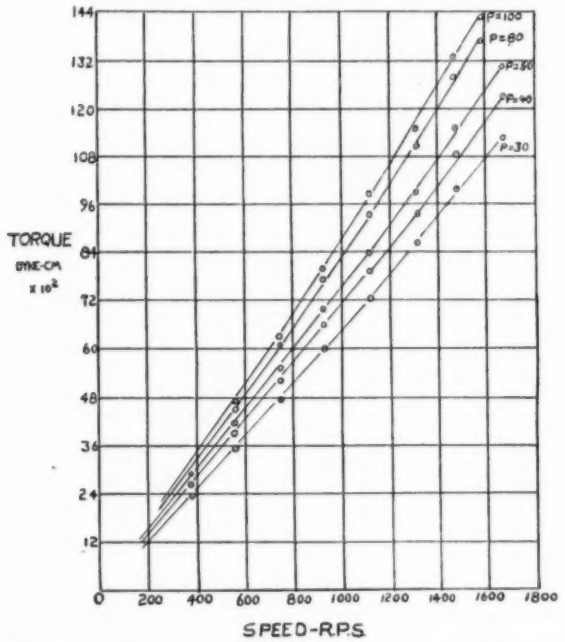



FIG. 7. Torque *vs.* rotor speed. $P$ = pressure in microns. Temperature = 26°C.

Figure 8 shows a plot of torque *versus* pressure, using the rotor speed as the parameter. The general form of the curves can be seen to be independent of the speed, and the curves appear to be asymptotic to a certain torque value. Several observations were also made at pressures of 150 and 200 microns, and it may be stated that, within the limits of experimental accuracy, the results obtained agreed with the general pattern of Figs. 7 and 8.

## DISCUSSION

The fact that the curves of Fig. 8 exhibit the property of approaching some torque value asymptotically indicates that the retarding torque on the rotor may be due to ordinary gaseous viscosity, since one would expect the coefficient of viscosity, and hence the retarding torque, to be almost independent of the pressure at ordinary pressures. Some time ago, Itterbeek and Paemel[4] measured the coefficient of viscosity of several pure gases as a function of pressure at room temperatures using the oscillating disk method. They found that for all gases used, the results could be expressed by the theoretically derivable equation $\eta = (\eta_\infty p / a\eta_\infty + p)$ where $\eta$ is the coefficient of viscosity at the pressure, $p$; $\eta_\infty$ is the coefficient of viscosity at atmospheric pressure, and $a$ is a parameter which is independent of the pressure. Attempts to reconcile the present observations with this equation met with partial success. A fairly constant value of the parameter, $a$, was obtained in the pressure range from 1 to 10 microns, but at higher pressures the constancy failed. This narrow pressure interval corresponds to a region where, for the apparatus used, the intermolecular collision frequency is of the same order of magnitude as that of the molecules with the walls and is generally recognized to be below the so-called "slip" region. As the pressure was increased above 10

[4] A. van Itterbeek and O. van Paemel, Physica **7**, 273 (1940).

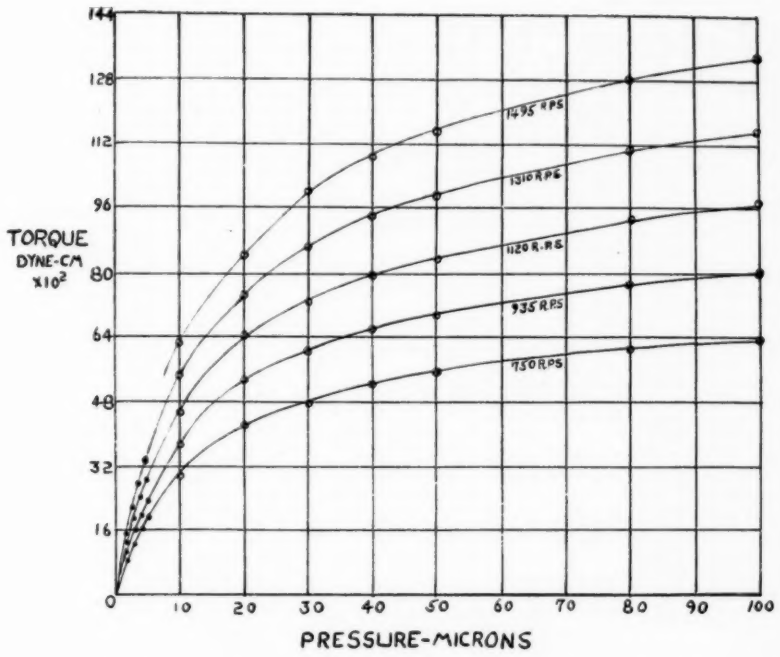



FIG. 8. Torque *vs.* pressure. Temperature = 26°C.

microns, the value of the parameter consistently increased. This meant that the values of $\eta$ as calculated from the experimental results were becoming smaller than those predicted by the theory. This fact was substantiated by a comparison of these experimental values of $\eta$ with values obtained by other investigators using very low speed viscosimeters. For example, Stacy[5] obtained the value of approximately $16\times10^{-5}$ c.g.s. units as 0.1 mm Hg, while the present observations yield $11\times10^{-5}$ c.g.s. units at this pressure.

The author wishes to record his indebtedness to Prof. J. W. Beams, upon whose suggestion this work was undertaken, and to various other members of the staff of the Rouss Physical Laboratory for their helpful advice and criticism.

[5] L. J. Stacy, Phys. Rev. **21**, 293 (1923).

# Letters to the Editor

## On the Flow of a Perfect Fluid between Two Plates from a Circular Inlet Channel

W. R. SMYTHE
*California Institute of Technology, Pasadena, California*
December 3, 1948

IN a Letter to the Editor in the November 1948 Journal of Applied Physics, p. 1092, E. T. Benedikt gives a definite integral formula for the velocity potential $\phi$ of an ideal incompressible, non-viscous fluid flowing from the region between two thick parallel plates a distance $h$ apart into a circular cylindrical hole of radius $r_0$ in one of them. For some purposes it may be more convenient to express his result as a series. This could be done by contour integration, but it is shorter to start over. The velocity potential $d\phi$ of a ring source of linear strength $2qdr'$ at $z=0$, $r=r'$ between two plates at $z=h$ and $z=-h$ is found by the usual method[1] to be, when $r<r'$,

$$d\phi_1=(qr'dr'/h)[\ln r'+2+\sum_{n=1}^{\infty} I_0(Nr)K_0(Nr')\cos Nz], \quad (1)$$

where $I_0(x)$ and $K_0(x)$ are modified Bessel Functions, and $N$ is $n\pi/h$. An interchange of $r$ and $r'$ inside the bracket gives the potential $d\phi_2$ when $r>r'$. For the uniform disk source of radius $r_0$ and strength $q$ per unit area assumed by Benedikt the potential is, according as $0<r<r_0$ or $r_0<r$,

$$\phi_1=\int_{r=0}^{r=r} d\phi_2+\int_{r=r}^{r=r_0} d\phi_1 \quad \text{or} \quad \phi_2=\int_0^{r_0} d\phi_2. \quad (2)$$

The velocities $v_r$ and $v_z$ are the quantities of chief interest. Integration for $\phi_1$ and $\phi_2$, use of the relation $v^{-1}=I_1(v)K_0(v)+I_0(v)K_1(v)$, and differentiation with respect to $r$ give $v_r$ when $r<r_0$ to be

$$(qr_0/h)[\tfrac{1}{2}(r/r_0)-2\sum_{n=1}^{\infty} I_1(Nr)K_1(Nr_0)\cos Nz].$$

When $r>r_0$, interchange $r$ and $r_0$ inside the bracket. Differentiation with respect to $z$ gives for $v_z$

$$r<r_0, \quad (q/h)[h-z-2r_0\sum_{n=1}^{\infty} I_0(Nr)K_1(Nr_0)\sin Nz],$$

$$r>r_0, \quad (2qr_0/h)\sum_{n=1}^{\infty} I_1(Nr_0)K_0(Nr)\sin Nz. \quad (3)$$

Except when $r=r_0$, the convergence is very rapid because the $I_n(v)K_n(v)$ product becomes proportional to $\exp(-n\pi|r-r_0|/h)$ as $n\to\infty$.

As Benedikt points out, the assumption of a uniform source gets rapidly worse as $h$ decreases. If $h\ll r_0$, however, another method gives increasingly accurate results. The distribution of flow across the gap between the plates then approaches that of the two-dimensional flow into a crack of width $2h$ in an infinite plane faced block. At $r=r_0$ the radial velocity $v_r$ is then given[2] to one part in 5000 by the formula

$$v_0\{A+B\cos(\pi z'/h)+C(z'/h)^2+D[\sec(\tfrac{1}{2}\pi z'/h)]^{0.355}\}, \quad (4)$$

where $z'$ is measured from the unperforated plate $0<z'<h$ and $A$, $B$, $C$, and $D$ are numerical constants.[2] The mean radial velocity at $r=r_0$ is $v_0$. When $r>r_0$, $v_r$ may be written as the series

$$v_0\{1+\sum_{n=1}^{\infty} A_n[K_1(Nr)/K_1(Nr_0)]\cos(Nz)\}. \quad (5)$$

Where the coefficients $A_n$ have been calculated[2] for $0<n<21$. When $r<r_0$, the boundary conditions are $v_z=0$ at $z'=0$, $v_r=0$ when $h<z'<\infty$, and $v_r$ is given by (4) when $0<z'<h$. If the velocity is assumed uniform in the tube at some value of $z'$ such as $z'=5h$, the velocity potential when $0<z'<5h$ can be expressed as a Fourier series somewhat similar to (5) whose coefficients involve gamma-functions. Otherwise a Fourier integral is required.

[1] W. R. Smythe, *Static and Dynamic Electricity* (McGraw-Hill Book Company, Inc., New York, 1939), Art. 5.324.
[2] W. R. Smythe, Rev. Mod. Phys. **20**, 176 (1948).

## Small Spherical Particles of Exceptionally Uniform Size

ROBERT C. BACKUS AND ROBLEY C. WILLIAMS
*Department of Physics, University of Michigan, Ann Arbor, Michigan*
November 30, 1948

DURING the course of some quantitative work with the electron microscope we have chanced upon some small, spherical particles of remarkable uniformity of size. The particles are contained in a polystyrene latex, made by the Dow Chemical Company, and are described as Dow latex 580-G, Lot 3584. Their dimensional uniformity immediately suggests a number of applications in electron microscopy, and three of the uses will be discussed here.

1. *Magnification Calibration*:—Inasmuch as the particles are readily suspended in distilled water, they can be applied directly with a small pipette to an electron microscope specimen, and can serve as a standard of size for any pictures taken of that specimen. An example of the appearance of the latex, diluted 1000 to 1 with water, deposited upon a specimen, is shown in Fig. 1. In this case the speci-

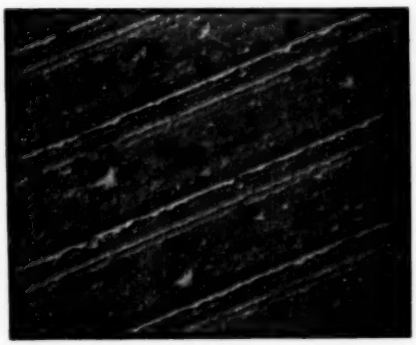


FIG. 1. Electron micrograph of Dow latex polystyrene particles supported upon a pre-shadowed replica of a 15,000-line/inch diffraction grating.

men is a shadow-cast collodian replica of a glass diffraction grating of 15,000 lines per inch.

A determination of the mean diameter of the individual particles has been made, as well as a determination of the dispersion in diameter from particle to particle. The method first tried for a determination of size was to prepare shadow-cast, organic replicas of a diffraction grating of known spacing, and then to make a direct comparison with the particle diameter, as could be done by the use of Fig. 1. After a few measurements, however, it became apparent that the replicas were more dimensionally variable than were the particles. We then decided to make a determination of particle size through the intermediate use of a light microscope whose eyepiece reticle had been calibrated against a standard etched-glass scale. The first step in the measurement was to allow the objects to adhere to the surface of fine glass fibers, following immersion of the fibers in a concentrated suspension of the particles. The partially coated fibers were then spread at random across a microscope specimen screen. Upon examination with the light microscope certain unique clusters of particles could be identified, some 40 to 50 microns apart, and their separation could be measured with the use of the calibrated reticle. The same fiber was then found in the field of the electron microscope, and by the use of several overlapping exposures, the distance along the fiber between the clusters could be measured on the micrographs. These measurements, of course, allowed a determination to be made of the diameters of the numerous individual particles found between the clusters. It was thought that a correction for distortion in the electron micrographs would be necessary, but no measurable difference in magnification exists across the field of the micrographs. The instrument used is an RCA, type EMB, of an early make, and the pictures were taken at a magnification of about 10,000.

The result of the measurements is a value of $2590 \pm 25$A for the mean diameter of the particles.* The probable error is thought to reside largely in the errors of measurement, rather than in a dispersion of diameter of the objects. Observations of the appearance of the shadow-cast particles, as well as of the appearance of the particles when seen in silhouette along the glass fibers, lead us to believe that the objects are truly spherical.

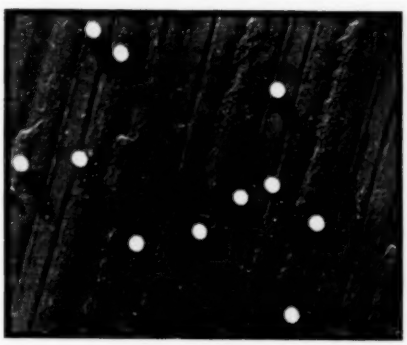


FIG. 2. Electron micrograph of uranium-shadowed Dow latex particles supported upon a replica of a diffraction grating.

A determination has also been made of the dispersion in diameter of about 500 of the small spheres, photographed on 20 pictures of identical magnification. The values of the diameters yield an internal probable error of $\pm 10$A, indicating a very high degree of dimensional uniformity. In measuring this large number of particles, we found there are some (about 2 percent of the total) which are clearly "sports" in that their diameters are different from the ordinary ones by almost a factor of two.

2. *Contours of Replica Surfaces*:—A second use of the latex spheres is illustrated by Fig. 2, which is a micrograph of the particles as placed upon the specimen film followed by uranium shadowing of particles and substrate. The ratio of length of shadow to diameter of sphere gives the value of the local angle of shadowing, and since this is frequently not constant over the area of a specimen, its evaluation at any one point is desirable. By comparing the lengths and shapes of the shadows of the particles over an extended region of the specimen, one may obtain a fair quantitative notion of the contours of the surface. The shadow-casting technique itself, of course, gives one a qualitative impression of specimen contours, but the shadows of the spherical particles allow a more quantitative estimate to be made. In Fig. 2, for example, one may note the considerable difference in shadow lengths, depending upon whether the particle is on an "up-hill" or "down-hill" portion of the replica surface.

3. *Thickness of Film Used in Shadowing*:—If some latex particles are placed on the specimen prior to shadowing, and some subsequent to shadowing, evidence can be obtained of the thickness of the evaporated film used for shadowing. It will be noticed in Fig. 2 that the spheres are slightly elongated in the direction of the length of the shadows. The elongation is due to the accretion of the evaporated film on the side of the particle nearest the filament during the shadow-casting; in Fig. 2 the elongation is about 35A. The thickness of the evaporated film on the surface of the substrate is calculated by measuring the difference between the mean diameter of the unshadowed particles and the mean major diameter of the shadowed ones, and dividing the difference in diameters by the ratio of the shadow length to the particle diameter.

* This is 50A greater than the preliminary value reported in the abstracts of the E. F. Burton Memorial meeting of the Electron Microscope Society of America.

## The Coefficient of Anomalous Viscosity

CHARLES MACK
*Technical and Research Department, Imperial Oil, Ltd., Sarnia, Ontario*
December 1, 1948

IN a recent paper on "Anomalous flow as an order-disorder transition" the writer gave the following expression for the coefficient of viscosity for anomalous flow:[1]

$$\eta = (n_\beta - n_\alpha) b\sigma / v, \qquad (1)$$

where $\sigma$=shearing stress, $v$=shear rate, $n_\beta$=degree of disorder, $n_\alpha$=degree of order, $b$=constant equal to $1-\alpha$ with $\alpha$ being the ratio of the number of holes to the number of molecules. In the subsequent derivation of $n_\beta - n_\alpha$ as a function of $\sigma$ and $v$, the viscosity was considered to be constant. It can be shown that this treatment holds only if the viscosity is expressed as $\eta_0$, the viscosity at a stress $\sigma_0$ near the yield value.

Since a change in volume is postulated, the actual viscosity decreases with increasing volume in the same manner as for a Newtonian liquid with increasing temperature, and therefore volume. For the latter, Batschinsky[2] gave the following equation:

$$\eta = C/(V-B), \tag{2}$$

where $V$=specific volume, and $C$ and $B$ are constants. Constant $B$ is of the same order of magnitude as the covolume in van der Waal's equation, and the term $V-B$ becomes consequently equal to the free volume of the liquid or the volume of the holes. This equation can also be applied to anomalous flow. If $N_h=\alpha N$ holes have a volume of $V_h$, the volume of one hole is $V_h/\alpha N$. With $\alpha N_\beta$ holes formed on the $\beta$-sites and $\alpha N_\alpha$ holes closed on the $\alpha$-sites, the free volume at a given stress becomes $\alpha(N_\beta - N_\alpha)Vh/\alpha N=(n_\beta-n_\alpha)V_h$, and Eq. (2) in combination with Eq. (1) may be written

$$\frac{\eta}{\eta_0}=\frac{(n_{\beta_0}-n_{\alpha_0})V_{h0}}{(n_\beta-n_\alpha)V_h}=\frac{(n_\beta-n_\alpha)b\sigma v_0}{(n_{\beta_0}-n_{\alpha_0})b\sigma_0 v},$$

and

$$n_\beta-n_\alpha=(n_{\beta_0}-n_{\alpha_0})(V_{h0}\sigma_0 v/V_h\sigma v_0)^{\frac{1}{2}}. \tag{3}$$

With $v/v_0=(\sigma/\sigma_0)^{1/b}$, for example, $n_\beta-n_\alpha$ becomes

$$n_\beta-n_\alpha=(n_{\beta_0}-n_{\alpha_0})(V_{h0}/V_h)^{\frac{1}{2}}(\sigma/\sigma_0)^{(1-b)/2b}, \tag{4}$$

and

$$\eta=(n_{\beta_0}-n_{\alpha_0})(V_{h0}/V_h)^{\frac{1}{2}}(\sigma/\sigma_0)^{(1-b)/2b}b\sigma/v. \tag{5}$$

In most cases the increase in the volume of a hole is small enough, so that $(V_{h0}/V_h)^{\frac{1}{2}}$ approaches unity and can be omitted without introducing a serious error. The equations in connection with stress-, strain-, and time-hardening and thixotropy given in the paper change correspondingly.

In a similar way the modulus of rigidity for anomalous elasticity is

$$G=n_\beta b\sigma/\gamma. \tag{6}$$

Whereas the coefficient of anomalous viscosity is a function of the free volume, the moduli of anomalous elasticity are independent of volume changes at constant temperature. It was shown in the paper, that in the case of polychloroprene the shear strain as a function of the shearing stress is $\gamma/\gamma_0=(\sigma/\sigma_0)^{1/b}$ with $b=1+\alpha$, hence,

$$n_\beta=n_{\beta_0}(\sigma_0/\sigma)^{(b-1)/b}. \tag{7}$$

With $b$ being constant, the shear rate becomes always a function of $\sigma^{1/b}$ for the various types of anomalous flow. Equation (1) holds also for cases where $b=1-\alpha$ is variable. Such a case is obtainable from the theory of von Mises and Hencky[3] which requires that for the condition of plasticity the product of the energy due to a change in shape and the modulus of rigidity should be constant—

$$WG=\text{constant}^2. \tag{8}$$

For the case of simple shear $WG$ is equal to $\sigma^2$. Applying this condition to plastic flow, the effective shearing stress per element of flow becomes $(1-\alpha)\sigma$, hence,

$$[(1-\alpha)\sigma]^2=[(1-\alpha_0)\sigma_0]^2, \tag{9}$$

indicating that the number of holes increases with increasing shearing stress. The coefficient of viscosity can be expressed by $\eta=d\sigma/dv_i$, where $v_i$ is the shear rate of an individual element of flow. It has been shown[4] that $dv_i=dv/(n_\beta-n_\alpha)$. Introducing this term into Eq. (1) and combining with Eq. (9) gives

$$\eta=(n_\beta-n_\alpha)d\sigma/dv$$
$$=(n_\beta-n_\alpha)(1-\alpha)\sigma/v=(n_\beta-n_\alpha)(1-\alpha_0)\sigma_0/v,$$

which on integration between $\sigma$ and $\sigma_0$, $v$ and $v_0$ leads to

$$\exp[(\sigma-\sigma_0)/\sigma_0]=(v/v_0)^b, \tag{10}$$

the Prandtl equation (Eq. (29) in the paper) with $b=1-\alpha_0$ =constant.

[1] C. Mack, J. App. Phys. **19**, 1082 (1948).
[2] A. J. Batschinsky, Zeits. f. physik. Chemie **84**, 643 (1913).
[3] A. Nadai, *Plasticity* (McGraw-Hill Book Company, Inc., New York, 1931), p. 70.
[4] C. Mack, J. App. Phys. **17**, 1087 (1946).

## Measurement of Gain of Electromagnetic Horns

ALLEN S. DUNBAR AND M. DONALD ADCOCK
*Antenna Research Section, Naval Research Laboratory, Washington, D. C.*
November 17, 1948

REFERENCE is made to "Gains and effective areas of horn antennas," by R. B. Watson and C. M. McKinney.[1] The method for measuring gain and the experimental technique described lead to large errors and fail to point out the relation of the observed gain of an antenna to the theoretical gain. The following method is precise and leads to results consistent with the gain factors discussed below.

Given two antennas separated by a large distance $R$ with power $P_T$ being delivered to the transmitting antenna, the power $P_R$ received by the second antenna is

$$P_R=\frac{P_T}{4\pi}\frac{G_T A}{R^2}=\frac{P_T G_T G_R \lambda^2}{(4\pi)^2 R^2}, \tag{1}$$

where $G_T$ and $G_R$ are the absolute gains of the transmitting and receiving antennas, respectively, and $A$ is the aperture area of the receiving antenna. If two identical matched antennas are used, and if their gains are assumed identical, then

$$G=(4\pi R/\lambda)(P_R/P_T)^{\frac{1}{2}}. \tag{2}$$

The two antennas are separated by a distance $R$ greater than $2D^2/\lambda$, where $D$ is the width of the aperture, in order to obtain a uniform phase and to avoid appreciable interaction. Thus $G$ may be determined by measuring the ratio of $P_R/P_T$ instead of $P_R$ and $P_T$ separately. This ratio may be determined easily within 2 percent. The wave-length and the distance $R$ may be measured with good accuracy.

This assumption that the two antennas have identical gains may be verified by using the same matched bolometer detector with each of them to see that the power received from a suitable third antenna is the same for both. If the gains are different, a proportionality factor may be introduced, and the true gains computed from (2). This procedure then produces two absolute gain standards at the operating wave-length.

The gain $G_x$ of any other antenna may be determined by using the same matched bolometer detector with it and with one of the standards to receive power from a third transmitting antenna. Then

$$G_x=(P_x/P_s)G, \tag{3}$$

where $P_x/P_s$ is the ratio of the power received by the unknown antenna to the power received by the gain standard.[2]

It is possible to calculate the gain of both conical and pyramidal horns by integration of the Poynting vector over the aperture.[3] This procedure leads to the following results.

$$G_{\text{conical}}=10.5 \text{ area of aperture}/\lambda^2, \tag{4}$$

$$G_{\text{pyramidal}}=10.2 \text{ area of aperture}/\lambda^2. \tag{5}$$

Hence the observed gain of a given horn is less than the so-called theoretical gain, defined by

$$G=4\pi \text{ area of aperture}/\lambda^2. \tag{6}$$

The ratio of the observed to the theoretical gain defines a "gain factor," which in the case of conical and pyramidal horns is 0.83 and 0.81, respectively.

[1] R. B. Watson and C. M. McKinney, J. App. Phys. **19**, 871–876 (1948).
[2] S. Silver, *Microwave Antenna Theory and Design* (McGraw-Hill Book Company, Inc., New York) (in press) Chapter 15; M. Katzin, "Quantitative radar measurements," Proc. I.R.E. **35**, 1333–1334 (1947).
[3] See reference 2, Chapter 10.

## A Method of Measuring Dynamic Mechanical Properties of Rubber-Like Materials

K. W. Hillier, H. Kolsky, and J. P. Lewis
*Imperial Chemical Industries, Butterwick Research Laboratories, Welwyn, Herts, England*
October 25, 1948

THE paper by A. W. Nolle[1] was read here with great interest. For the past three years we have been carrying out experiments using a technique similar to the method of longitudinal oscillations described on page 763 in that paper. Our preliminary results are in publication.[2] For our experiments, a more exact equation of propagation, assuming that reflection occurs at the pick-up, was derived. This reflection can be shown to affect both phase and amplitude measurements.

Assuming propagation at a velocity $V$, attenuation in the material to be proportional to $e^{-\alpha l}$ and the reflection coefficient to be $m(<1)$, independent of amplitude, then it can be shown that if $A\sin\omega t$ is the displacement at the source and $B\sin(\omega t+\beta)$ is the displacement at the pick-up

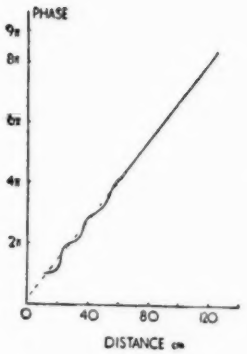



FIG. 1.

distance $l$ away: then

$$\tan\beta=\frac{1+me^{-2\alpha l}}{1-me^{-2\alpha l}}\tan kl, \tag{i}$$

and

$$B/A=(1-m)e^{-\alpha l}(1-2me^{-2\alpha l}\cos 2kl+m^2e^{-4\alpha l})^{-\frac{1}{2}}. \tag{ii}$$

For large values of $l$ and within a wide range of values of $\alpha$, Eq. (*i*) reduces to $\beta=kl$ as can be seen from a typical result shown in Fig. 1. Similarly for $l$ and $\alpha$ large, $B/A=Ke^{-\alpha l}$, and hence a plot of the logarithm of amplitude against distance gives values of $\alpha$ similar to the method used by Nolle. Our method yields results over a wider range of constants and by enabling us to use a continuous phase/distance relation we can obtain better values for the wavelength of the sound and hence for the velocity of propagation in the material.

[1] A. W. Nolle, J. App. Phys. **19**, 753 (1948).
[2] K. W. Hillier and H. Kolsky, Proc. Phys. Soc. (in press).

## Comment on "A Method of Measuring Dynamic Mechanical Properties of Rubber-Like Materials"

[J. App. Phys. **19**, 753 (1948)]
A. W. Nolle
*Department of Physics, University of Texas, Austin, Texas*
November 9, 1948

IT is true that by making a correction for reflection from the receiving element one can sometimes extend the useful range of the experiment in question to lower frequencies. The effects of reflection are considered very extensively in the author's thesis,[1] in which the influence of reflection upon the phase record is considered much as in the note of Hillier, Kolsky, and Lewis, and in which an approximate scheme for correcting the amplitude record is derived.

The author's experience has been that in the case of most types of rubber-like materials the fractional extension of the useful ranges of frequency and temperature in the experiment, made possible by reflection correction, is relatively small. Further, making corrections for reflection represents a considerable increase in the effort required to

process the data, with the result that it is often preferable to make use of a different experimental method rather than to employ corrections in the constant-frequency technique for longitudinal-wave study. For these reasons discussion of the corrections appeared to constitute an unjustifiable addition to the length of the paper cited by the above authors and was therefore omitted.

[1] A. W. Nolle, *Dynamic Mechanical Properties of Rubber-Like Materials* (Ph.D. thesis), Massachusetts Institute of Technology, August, 1947.

*Note:* Later information which has reached the author concerning the work of Hillier, Kolsky, and Lewis indicates that much of their work has been at relatively low frequencies (1 to 2 kc) and has been concerned with polyethylene, which has less acoustic attenuation than do some of the usual "rubbers." Thus the views expressed in these two letters are in substantial agreement.

## An Extension of MacFarlane's Method of Deducing Refractive Index from Radio Observations

A. W. STRAITON
*The University of Texas, Austin, Texas*
October 29, 1948

MACFARLANE[1] has presented relationships whereby the refractive index may be determined for a stratified atmosphere from radio data. These equations apply at sufficient distance beyond the optical horizon so that only one mode of propagation is significant. Because of difficulties in measurement, phase was eliminated and the results presented in terms of magnitudes only.

Recent work at The University of Texas[2] and elsewhere has shown the feasibility of measuring phase of centimeter waves. If measurements of phase are admitted, considerable simplification in procedure is obtained.

The field equation solution may be expressed as the sum of a number of normal modes as

$$\psi = e^{i\omega t - i\pi/4}(2\pi/kR)^{\frac{1}{2}} \Sigma\, e^{-ikSR} U_m(h_t) U_m(h_r), \qquad (1)$$

where $R$ is the distance along the earth's surface, $K$ is $2\pi$ divided by wave-length, $S_m$ is the propagation function for the $M$ mode, $U_m(h_t)$ and $U_m(h_r)$ are the height-gain relationships for the transmitter and receiver, respectively, and $U_m$ is a solution of the equation

$$(d^2U_m/dh^2) + k^2(M^2 - S_m^2)U_m = 0, \qquad (2)$$

where $M$ is the index of refraction modified to correct for the earth's curvature and $S_m$ is the characteristic value of the $M$ mode.

Since $S_m$ and $U_m$ are complex, they may be replaced by

$$U_m = Ae^{i\phi_m} \quad \text{and} \quad S_m^2 = a + ib.$$

Substituting these into Eq. (2) gives

$$\ddot{A} - A(\dot{\phi})^2 = k^2(a - M^2)A \qquad (3)$$

$$2\dot{A}\dot{\phi} + A\phi = k^2 bA. \qquad (4)$$

MacFarlane eliminated $\phi$ from these equations and found

$$k^2(a - M^2) = \ddot{A}/A - (k^4b^2/A^4)\left(\int_0^h A^2 dh\right)^2. \qquad (5)$$

This gives a value of $a - M^2$ in terms of the magnitude of signal strength. Since $a$ is very nearly one, the modified index $M$ may be determined except for a small constant which is independent of height.

The solution in terms of signal strength only requires a complete height-gain function from the reference height to the height $h$. In addition, another height-gain curve at a different frequency or measurements at different distances are necessary in order to determine the attenuation factor $b$.

However, using phase measurement, only one set of height-gain and height-phase measurements is necessary, and measurements over a limited range only are required since no integration is necessary. From Eq. (4), the attenuation factor is given by

$$b = [2\dot{A}\dot{\phi} + A(\partial^2\phi/\partial l^2)]/(k^2A) \qquad (6)$$

and from Eq. (3), $a - M^2$ is given by

$$a - M^2 = [\ddot{A}/A - (\dot{\phi})^2]/k^2. \qquad (7)$$

In the measurement method used at The University of Texas, $\dot{\phi}$ and $A$ are measured directly. The determination of $b$ thus involves the measured quantities and their rate of change, while the determination of $(a - M^2)$ involves in addition the second derivative of $A$.

Preliminary application of Eq. (6) to data taken on a 32-mile overwater path showed general agreement between the value of $b$ calculated from Eq. (6) and that determined from the $M$ curve. Detailed study of this and future data will be made in light of Eqs. (6) and (7).

[1] G. G. MacFarlane, "A method of deducing the refractive-index profile of a stratified atmosphere from radio observations," *Meteorological Factors in Radio Wave Propagation*, published by The Physical Society, London.

[2] A. W. Straiton and J. R. Gerhardt, "Results of horizontal microwave angle-of-arrival measurements by the phase difference method," Proc. I.R.E. **36**, 916–922 (1948).

## Errata: Energy Losses of Sound Waves in Metals due to Scattering and Diffusion

[J. App. Phys. **19**, 940 (1948)]

W. P. MASON AND H. J. MCSKIMIN
*Bell Telephone Laboratory, Murray Hill, New Jersey*

IN connection with the evaluation of the average shear modulus and the scattering factor for shear waves it appears[1] that the direction cosines of Eqs. (12) are not sufficiently general to produce a true average of these quantities for all directions. The most general direction cosines are those given for a set of Euler coordinates:

$$\begin{aligned}
l_1 &= \cos\theta\cos\varphi\cos\psi - \sin\varphi\sin\psi;\\
l_2 &= -\cos\theta\cos\varphi\sin\psi - \sin\varphi\cos\psi;\\
l_3 &= \cos\varphi\sin\theta;\\
m_1 &= \cos\theta\sin\varphi\cos\psi + \cos\varphi\sin\psi;\\
m_2 &= \cos\varphi\cos\psi - \sin\varphi\sin\psi\cos\theta;\\
m_3 &= \sin\varphi\sin\theta;\\
n_1 &= -\sin\theta\cos\psi;\\
n_2 &= -\sin\theta\sin\psi;\\
n_3 &= \cos\theta.
\end{aligned}$$

Introducing these values into Eqs. (11) and (14), which give the values of $c_{44}'$ for any orientation for cubic and hexagonal crystals, respectively, one can obtain the average

values for the shear modulus by performing the integration

$$\langle c_{44}' \rangle_{Av} = \int_0^{2\Pi} \frac{d\psi}{2\Pi} \int_0^{2\Pi} \frac{d\varphi}{2\Pi} \int_0^{\Pi} \frac{c_{44}' \sin\theta d\theta}{2}.$$

For cubic crystals the average value becomes

$$\langle c_{44}' \rangle_{Av} = c_{44} + \frac{c_{11} - c_{12} - 2c_{44}}{5} = \frac{3}{5} c_{44} + \frac{c_{11} - c_{12}}{5}.$$

For hexagonal crystals, the result is

$$\langle c_{44}' \rangle_{Av} = \frac{7}{30} c_{11} - \frac{c_{12}}{6} - \frac{2}{15} c_{13} + \frac{c_{33}}{15} + \frac{2}{5} c_{44}.$$

The scattering factor for shear waves is likewise altered, and the correct value for cubic crystals is

$$\left\langle \left[ \frac{c_{44}' - \langle c_{44}' \rangle_{Av}}{\langle c_{44}' \rangle_{Av}} \right]^2 \right\rangle_{Av} = \frac{3}{175} \left[ \frac{c_{11} - c_{12} - 2c_{44}}{\langle c_{44}' \rangle_{Av}} \right]^2.$$

The numerical values corresponding to the average values of $\langle c_{44}' \rangle_{Av}$ and the scattering factors for Tables I, II, and III are given in the attached table.

| Material | $\langle c_{44}' \rangle_{Av}$ | $\left\langle \left[ \frac{c_{44}' - \langle c_{44}' \rangle_{Av}}{\langle c_{44}' \rangle_{Av}} \right]^2 \right\rangle_{Av}$ |
|---|---|---|
| Al | $2.62 \times 10^{11}$ dynes/cm² | $3.3 \times 10^{-2}$ |
| Au | 3.15 | $5.2 \times 10^{-2}$ |
| Ag | 3.21 | $6.1 \times 10^{-2}$ |
| Cu | 5.46 | $6.7 \times 10^{-2}$ |
| Pb | 1.016 | $7.2 \times 10^{-2}$ |
| Fe | 8.86 | $4.0 \times 10^{-2}$ |
| Ng | 0.384 | $1.25 \times 10^{-1}$ |
| K | 0.175 | $1.1 \times 10^{-1}$ |
| W | 15.15 | 0 |
| Mg | 1.77 | |
| Zn | 4.52 | |
| Cd | 2.55 | |

[1] Professor L. Gold, Brown University, private communication.

## A Note on the Ionospheric Absorption Problem

LESLIE G. McCRACKEN
*Electrical Engineering Department, The Pennsylvania State College, State College, Pennsylvania*
November 2, 1948

IN the well-known Bakerian Lecture delivered by Sir E. V. Appleton,[1] the total non-deviating $E$ layer absorption, under the condition that the frequency of the transmitted wave was much greater than the collisional frequency, was stated to be:

$$-\log\rho = 4.13 \frac{4\pi e^2}{mc} \frac{N_0 \nu_0 (\cos\chi)^{\frac{1}{2}}}{(p+p_L)^2}. \quad (1)$$

Here: $\rho$ = reflection coefficient, $e$ = electronic charge, $m$ = mass of electron, $c$ = velocity of light, $N_0$ = electron density at the maximum of the layer when $\chi = 0$, $\nu_0$ = collisional frequency at the height of $N_0$, $H$ = scale height, $\chi$ = zenith angle, $p$ = angular frequency of propagated wave, and $p_L$ = angular gyro-magnetic frequency due to the longitudinal component of earth's field.

Almost simultaneously, J. E. Best and J. A. Ratcliffe[2] derived a similar expression for the absorption, their expression being:

$$-\tfrac{1}{2}\log\rho = \frac{1.25}{c} \frac{4\pi e^2}{m(p+p_L)^2} H \nu_0 \left(\frac{B}{\alpha}\right)^{\frac{1}{2}} \left(\frac{\cos\chi}{C}\right)^{1.5}, \quad (2)$$

where $\nu_0$ in this relation refers to the value of $\nu$ at the ground. After proper substitutions for the constants $B$ and $\alpha$, the expression for the absorption becomes identical.

It is the purpose of this note to derive the expression for the absorption without using the approximations appearing in Best and Ratcliffe's paper, retaining, however, the condition that $p^2 \gg \nu^2$ in the region where the absorption is appreciable.

As a starting point, we select

$$-\tfrac{1}{2}\log\rho = \int_0^{\infty} k dh, \quad (3)$$

where $k$ = absorption coefficient as a function of height, and $h$ = height above the earth's surface. Upon substituting the relation for the absorption coefficient as a function of height, Best and Ratcliffe set up the equation above as

$$-\tfrac{1}{2}\log\rho = \frac{1}{2mc} \frac{4\pi e^2}{(p+p_L)^2} \nu_0 \left(\frac{B}{\alpha}\right)^{\frac{1}{2}} \times \int_0^{\infty} \exp\left\{ -\frac{3}{2}\frac{h}{H} - \frac{C\sec\chi}{2} \exp\left(-\frac{h}{H}\right) \right\} dh, \quad (4)$$

and then proceeded to evaluate the integral

$$\int_0^{\infty} \exp\left\{ -\frac{2}{3}\frac{h}{H} - \frac{C\sec\chi}{2} \exp\left(-\frac{h}{H}\right) \right\} dh.$$

They define $x = h/H$, so that the integral becomes

$$H \int_0^{\infty} \exp\left\{ -\tfrac{3}{2}x - \frac{C\sec\chi}{2} \exp(-x) \right\} dx.$$

By a change of variable, Best and Ratcliffe obtain a relation similar to Eq. (5) below. They then immediately go to an approximation. It is the primary purpose of this note to indicate the restrictions set up by their approximation. A further purpose is to indicate a method whereby these restrictions may be avoided.

Let us define $Z = \frac{C\sec\chi}{2} e^{-x}$, so that

$$dZ = -\frac{C\sec\chi}{2} e^{-x} dx,$$

$$dx = -\frac{2}{\sec\chi} e^{x} dZ = -\frac{dZ}{Z},$$

$$e^{-x} = \frac{2Z}{C\sec\chi}.$$

We find that the integral becomes

$$H \int_{C\sec\chi/2}^{0} \left(\frac{2Z}{C\sec\chi}\right)^{\frac{1}{2}} e^{-Z} \left(-\frac{dZ}{Z}\right) = H \left(\frac{2}{C\sec\chi}\right)^{\frac{1}{2}} \int_0^{C\sec\chi/2} Z^{\frac{1}{2}} e^{-Z} dZ. \quad (5)$$

The integral above is known as the incomplete factorial function and it may be expressed as

$$H \left(\frac{2}{C\sec\chi}\right)^{\frac{1}{2}} \left(\frac{1}{2}, \frac{C\sec\chi}{2}\right)!. \quad (6)$$

The factorial function may also be written in terms of the modified gamma-function for which tables are available.[3]

The integral finally may be stated as

$$H\left(\frac{2}{C\sec\chi}\right)^{\frac{1}{2}}\Gamma\left(\frac{3}{2},\frac{C\sec\chi}{2}\right), \tag{7}$$

whereupon the total absorption becomes

$$-\tfrac{1}{2}\log\rho=\frac{H}{2mc}\frac{4\pi e^2}{(p+p_L)^2}\nu_0\left(\frac{B}{\alpha}\right)^{\frac{1}{2}}\left(\frac{2}{C\sec\chi}\right)^{\frac{1}{2}}\Gamma\left(\frac{3}{2},\frac{C\sec\chi}{2}\right). \tag{8}$$

It is interesting to note certain features of the expression above. When the zenith angle is $\pi/2$, its secant becomes infinite, and the modified gamma-function reduces to the ordinary gamma-function. Thus, the total absorption for this case is

$$-\tfrac{1}{2}\log\rho=\frac{1}{2mc}\frac{4\pi e^2}{(p+p_L)^2}\nu_0 H\left(\frac{B}{\alpha}\right)^{\frac{1}{2}}\left(\frac{2}{C\sec\chi}\right)^{\frac{1}{2}}\frac{(\pi)^{\frac{1}{2}}}{2} \tag{9}$$

$$=\frac{1.25}{c}\frac{4\pi e^2}{m(p+p_L)^2}H\nu_0\left(\frac{B}{\alpha}\right)^{\frac{1}{2}}\left(\frac{\cos\chi}{C}\right)^{1.5}. \tag{10}$$

Equation (1) is identically equal to the results indicated by Appleton, Best, and Ratcliffe. Where the zenith angle approaches zero, the absorption must be evaluated with the modified gamma-function tables. As close as we are able to use the figure appearing in Johnke and Emde's *Tables of Functions*,[4] we find that the modified gamma-function for $\Gamma(\frac{3}{2},\frac{1}{2})=0.25\Gamma(\frac{3}{2},\infty)$, so that the total absorption when $\chi=0$ is

$$-\tfrac{1}{2}\log\rho=\frac{0.31}{c}\frac{4\pi e^2}{m(p+p_L)^2}\nu_0 H\left(\frac{B}{\alpha}\right)^{\frac{1}{2}}\left(\frac{\cos\chi}{C}\right)^{1.5}. \tag{11}$$

It would appear from this discussion that the absorption should still vary approximately as the three-halves power of the zenith angle's cosine, but that the relations given by Appleton, Best, and Ratcliffe are valid only for large zenith angles. This considerably restricts the value of their approximation since, as is well known, the Chapman[5] relation which is utilized in setting up Eqs. (1) and (2) is not valid for large zenith angles.

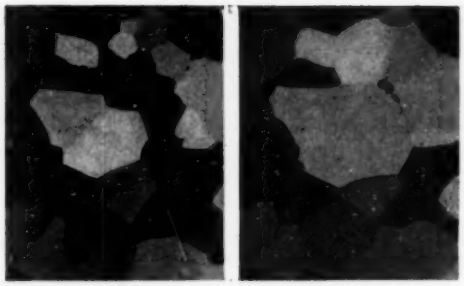
[1] E. V. Appleton, "The Bakerian Lecture, regularities and irregularities in the ionosphere—I," Proc. Roy. Soc. (London) **A162**, 451–479 (1937).
[2] J. E. Best and J. A. Ratcliffe, "The diurnal variation of the ionospheric absorption of wireless waves," Proc. Phys. Soc. **50**, 233 (1938).
[3] K. Pearson, *Tables of the Incomplete Gamma Function* (London, 1922).
[4] E. Johnke and F. Emde, *Tables of Functions* (Dover Publications, New York, 1945), fourth edition, p. 23.
[5] S. Chapman, "The absorption and dissociation or ionizing effect of monochromatic radiation in an atmosphere on a rotating earth. Part I," Proc. Phys. Soc. **43**, 26–45 (1931).

## Grain Growth in Octachloropropane

W. C. McCrone and P. T. Cheng
*Armour Research Foundation, Chicago, Illinois*
October 19, 1948

Beck[1,2] has shown that isothermal grain growth in a pure metal and in certain pure solid solutions follows the equation

$$D=K\cdot(tg+A)^n, \tag{1}$$



(a) (b)

Fig. 1. "Annealing" of octachloropropane; photomicrographs a and b were taken of the same field two hours apart while maintaining the temperature at 115°C.

where $tg$ is that portion of the total annealing time, $t$, which is available for grain growth, i.e., the total annealing time less the time, $R$, necessary for complete recrystallization: $tg=t-R$. The parameters $n$ and $K$ are functions of the temperature. The parameter $A$ is interpreted as the period of imaginary grain growth which would increase the grain size from zero to the recrystallized grain size $D_r$ in accordance with the extrapolation of the formula describing grain growth subsequent to recrystallization, e.g. (1).

In a recent note,[3] Beck proposed an assumption that the isothermal grain growth in specimens of the same material, starting from different recrystallized grain sizes, $D_r$, has the same instantaneous rate $dD/dt$ at corresponding grain size levels, and he is continuing work to prove this assumption by studying grain growth in brass after recrystallization to various grain sizes.

Our work on boundary migration using octachloropropane has given data on grain growth which, if judged applicable, fully confirms Beck's assumptions. The use of organic compounds to exemplify the behavior of metallic systems has been common practice, especially in teaching metallography. The advantages are readily apparent—lower temperature range, transparent specimens, as well as generally simpler techniques. The question, of course, remains as to the extent to which the crystallization behavior is analogous with metals systems. In the present case it seems probable that octachloropropane can be used to study grain growth in metals. The octachloropropane used was purified by sublimation from material furnished by the Hooker Electrochemical Company and advertised as $C_3H_{0.16}Cl_{7.84}$. The crystals are uniaxial with low birefrigence and give grains very similar in appearance to aluminum or, except for twinning, to brass. On warming to about 100°C (annealing) the grains grow (recrystallize) in the same manner as these two metals. Figure 1 shows the same field of octachloropropane taken about two hours apart while being held at 125°C. The data presented below were made on preparations made by fusing a few milligrams of octachloropropane between cover glass and slide. The melt on supercooling gives crystallization in the form of small grains with various orientations as shown by crossed nicols. Photomicrographs are then taken at regular intervals while the preparation is held at a uniform temperature in a hotstage. Grain size data are obtained from as many

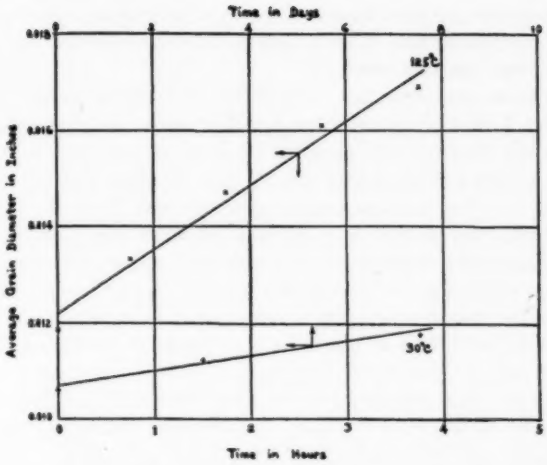



FIG. 2. Isothermal grain growth curves for octachloropropane.

grains as possible in each photomicrograph. These data are then handled as described by Beck.[1-3]

The results, shown graphically in Fig. 2, prove that grain growth in octachloropropane follows the relation suggested by Beck, $D=K\cdot(tg+A)^n$, and that $n$ is approximately one. Octachloropropane is then an interesting and useful system to study; the results may be helpful in leading to a better understanding of grain growth in metals.

[1] Beck, Kremer, Demer, and Holzworth, "Grain growth in high purity aluminum and in aluminum-magnesium alloy," AIME Tech. Pub. No. 2280 Metals Tech. (September 1947).
[2] Paul A. Beck, John Towers, Jr., and William D. Manly, "Grain growth in 70-30 brass," AIME Tech. Pub. No. 2326, Metals Tech. (February 1948).
[3] Paul A. Beck, "Effect of recrystallized grain size on grain growth," J. App. Phys. 19, 507 (1948).

## Comments on "Grain Growth in Octochloropropane"

[J. App. Phys. 20, 230 (1949)]

PAUL A. BECK
*Department of Metallurgy, University of Notre Dame, Notre Dame, Indiana*

October 27, 1948

DOCTORS McCrone and Cheng contributed an interesting new technique to the study of grain growth, and they found an organic material, the grain growth behavior of which appears to be very similar to that of metals. The beautiful photomicrographs in Fig. 1 show how suitable the new technique is for studying the details of the mechanism of grain growth.

The question of the numerical value of the parameter $n$ deserves further consideration. The data, even at 125°C, fall within a rather small grain size range. Consequently, it would require extraordinarily accurate grain size determinations to decide on a particular value for $n$. I understand the authors' tendency to draw a straight line and consider the deviations as a result of scatter. However, it may be observed that these deviations have a certain systematic character. It would appear that the points might satisfy a curve better than a straight line. They show a tendency of decreasing growth rate with increasing grain size, as in metals. If the results be plotted on a $\log D-\log t$ scale, it might easily happen that they would tend to support an $n$ value considerably lower than *one*. A more definite conclusion would have to be deferred, of course, until data were obtained covering a wider range of $D$ and $t$. I would be very much surprised if the $n$ value turned out to be larger than $\frac{1}{2}$.

*Additional comment by McCrone and Cheng (received November 10, 1948)*: Dr. Beck has correctly analyzed the situation with respect to the data presented. These data do seem to deviate systematically and consistently from a straight line and may indeed indicate an $n$ value less than one. Additional data covering a wider range of $D$ and $t$ are now being obtained, and an answer to this question should be available shortly. An effort will also be made to extend this technique to other similar compounds, both organic and inorganic.

## Comments on "Electronic Radiography and Microradiography"

[J. App. Phys. 19, 844 (1948)]

HERMAN E. SEEMANN
*Kodak Research Laboratories, Rochester, New York*

November 4, 1948

IN this paper Trillat describes certain experiments which may lead to useful metallurgical applications. Since their success depends to a great extent upon the choice and correct use of photographic materials, it seems appropriate to add a few supplementary remarks.

The principle involved is that the magnitude of the x-ray photoelectric effect depends upon the element irradiated, and consequently, the surface of an alloy, for example, may be studied by recording its electronic emission on a film which is in *intimate* contact with the surface. *Radiographs* of thin materials may be made using the electron emission from a *uniform* surface of high atomic number, such as lead. In either case, details are examined after enlargement. In general, emulsions which exhibit little graininess and are, therefore, best suited for enlargements of many diameters are slow; that is, they require relatively long exposures. With the wide variety of photographic materials available it should not be difficult to find some types that are well suited to most of the problems in electron and x-ray microradiography. However, I do not know of any film or plate capable of rendering a useful radiographic image when it is "magnified many thousands of times." Instead, the present-day limit seems to be a *few hundred times*.

At one point in the paper, the discussion of x-ray and electron sensitivity implies that a film is available which is sensitive to electrons only, but elsewhere it is brought out correctly that sensitivity to both electrons and x-rays exists in the one film. I do not know of any photographic emulsion which is sensitive to one and not to the other. It is to be expected that the radiation with the greater

absorption coefficient will have the greater effect. Hence, in "electronic radiography by the reflection method," the incident x-rays will always be present as a contrast-reducing factor in the radiograph, although the effect may be negligible in many cases.

Trillat recommends the use of "a vacuum apparatus when necessary" for insuring good contact between specimen and film. This point deserves considerable emphasis since what appears to be good contact at low magnification may be a very disturbing factor at high magnification. A vacuum cassette designed to insure good contact has been described by H. F. Sherwood.[1]

It would be better to say, in referring to Table I of Trillat's paper, that the photographic density increases with atomic number rather than that the photographic density "is exactly proportional to the atomic number." The relation between atomic number and density will depend upon the film used, its development, and the range of densities covered. In such experiments, *densities* should be converted to photographic *intensities* by means of the characteristic curve of the film being used. Thus, the characteristic curve is a calibration curve for the film as a measuring medium. (Under very *limited conditions* the relation between density and intensity may be linear.) Unfortunately, for ferrous metallurgical applications, the distinction between iron and its close neighbors is weak. A fine-grained film of very high contrast is indicated here to obtain the best results.

The results of certain experiments reported or referred to by Trillat corroborate some earlier work;[2] for example, I found that the photographic intensity of emission from metal foils increased with atomic number (Fig. 2 of reference 2). Also, a photographic negative was "copied" by utilizing the photoelectric emission from the silver in the original when it was irradiated with hard x-rays. The areal concentration of silver in a photographic image is so small that the emission depends largely upon the mass per unit area (p. 839). It was found that an "electron radiograph" of a piece of paper could be made by placing the sample between a film and a lead screen and exposing the combination to hard x-rays (p. 844). To estimate the energy of electrons emitted by lead foil with x-ray excitation, electron absorption curves were obtained for aluminum and Kodapak. In this way it was shown incidentally that the *mass* absorption coefficient was the same for both of these light materials (Fig. 4).

[1] H. F. Sherwood, "A vacuum exposure holder for microradiography," Rev. Sci. Inst. **18**, 80 (1947).
[2] H. E. Seemann, "Some physical and radiographic properties of metallic intensifying screens," J. App. Phys. **8**, 836 (1937).

# New Booklets

The Baldwin Locomotive Works, Philadelphia 42, Pennsylvania, has published its Bulletin 291, *The ABC-C-C of Simulated Service Testing.* It features simple cartoon sketches to emphasize the importance of knowing the effect of dynamic and cyclic stresses on many machine parts and structural members. 8 pages, available on request.

Training Section, Naval Research Laboratory, Washington 20, D. C., has issued a 36-page bulletin entitled *Research Accrediting at the Naval Research Laboratory.* Numbered P-557, it is intended for promising graduate students who wish to undertake basic research in the physical sciences.

The National Supply Company, Pittsburgh, Pennsylvania, in response to requests from advance readers, has made available to the public copies of an attractive brochure entitled *The Story of The National Supply Company.* Originally intended only as a company history, it has been described by outside readers as "a fast-moving and easily understood history of the oil industry." 48 pages.

Proceedings of the Illinois Conference on Industrial Research, sponsored by Armour Research Foundation and the Illinois Manufacturers' Association at Illinois Institute of Technology, Chicago, Illinois, May 27–28, 1948, are now in printed form. Price $1 per copy. Order from Public Relations Office, Armour Research Foundation, 35 West 33rd Street, Chicago 16, Illinois.

Dow Corning Corporation, Midland, Michigan, has announced a new technical publication entitled *DC 200 Silicone Fluids, Dow Corning Silicone Notebook, Fluid Series No. 3.* 32 pages, available on request.

The Phosphor Bronze Corporation, 2200 Washington Avenue, Philadelphia 46, Pennsylvania, has published a new technical data book on *Elephant Brand Phosphor Bronze.* 28 pages, free on request.

Struthers-Dunn, Inc., 150 North 13th Street, Philadelphia 7, Pennsylvania, has issued a new catalog bulletin designed to simplify the selection and use of midget relays for almost any application. The bulletin, No. 2100, is available on request.

James G. Biddle Company, 1316 Arch Street, Philadelphia 7, Pennsylvania, has announced publication of its Bulletin 41-29 on *Jagabi 'Lubri-tact' Tubular Rheostats.* It describes the various styles of standard, screw-drive, double, graded, switchboard, metal-cased, and non-inductive rheostats. 12 pages, free on request. Bulletin 35-29 on *Speed Measuring Instruments,* 28 pages, is also available from this company.

Infra Insulation, Inc., 10 Murray Street, New York 7, New York, has published a 32-page manual by Alexander Schwartz, president of the company, on *Simplified Physics of Thermal Insulation.* It covers heat transfer, condensation, vermin, mold, fire, radiant heating, and other facts and factors involved in the study of insulation and insulating values. Copies free on request.

The National Bureau of Standards recently announced the following publications, all available from the Superintendent of Documents, U. S. Government Printing Office, Washington 25, D. C.:

Circular 470 *Precision Resistors and Their Measurement*, by James L. Thomas. 32 pages, 20 cents.

Circular 473 *Measurement of Radioactive Isotopes*, by Leon F. Curtiss. 12 pages, 5 cents.

Circular 460 *Publications of the National Bureau of Standards*, listing of all Bureau publications from 1901 to June 30, 1947. Brief abstracts for the publications issued from January 1, 1942, to June 30, 1947, are also included. 375 pages, 75 cents.

The United Nations Educational, Scientific, and Cultural Organization recently published an illustrated booklet, *Suggestions for Science Teachers in Devastated Countries*. The author is J. P. Stephenson, science master at the City of London School. He explains how science teaching can be commenced without the use of apparatus and then shows how equipment for experiments can be improvised from commonplace materials. 88 pages.

Research Laboratory of Electronics, Massachusetts Institute of Technology, Boston 39, Massachusetts, has published Technical Report No. 41 by R. M. Fano entitled *Theoretical Limitations on the Broadband Matching of Arbitrary Impedances*. Copies of the full report are available on request. Following is the abstract:

"This paper deals with the general problem of matching an arbitrary load impedance to a pure resistance by means of a reactive network. It consists primarily of a systematic study of the origin and nature of the theoretical limitations on the tolerance and bandwidth of match and of their dependence on the characteristics of the given load impedance. Necessary and sufficient conditions are derived for the physical realizability of a function of frequency representing the input reflection coefficient of a matching network terminated in a prescribed load impedance. These conditions of physical realizability are then transformed into a set of integral relations involving the logarithm of the magnitude of the reflection coefficient. Such relations are particularly suitable for the study of the limitations on the bandwidth and tolerance of match. Definite expressions for these quantities are obtained in special cases. The practical problem of approaching the optimum theoretical tolerance by means of a network with a finite number of elements is also considered. Design curves are provided for a particularly simple but very important type of load impedance. In addition, a very convenient method is presented for computing the values of the elements of the resulting matching network."

Sinclair Refining Company, New York 20, New York, has issued a 20-page brochure to mark the opening of the new Sinclair Research and Development Laboratories at Harvey, Illinois, October 1948. Its title is *A Symbol of Progress*. Amply illustrated with photographs of the new laboratories.

Pittsburgh Plate Glass Company, Fifth Avenue at Bellefield, Pittsburgh 13, Pennsylvania, has published a 72-page booklet entitled *Chlorine*. It is the second in a series of technical data books and is designed for use by operators, technicians, buyers, and executives. Write for copy on your business letterhead.

Leeds and Northrup Company, 4934 Stenton Avenue, Philadelphia 44, Pennsylvania, has announced a 32-page illustrated catalog, ND44(1), which covers, for the first time in one publication, the company's complete line of recorders and controllers for measuring a wide variety of process variables. Free on request.

Industrial Distributors (1946) Ltd., 44 Main Street, Johannesburg, South Africa, have issued a 16-page brochure entitled *The Diamond Research Laboratory*. The brochure "presents to the technical world and the general public the aims and facilities of the Diamond Research Laboratory. This Laboratory has been established to assist all branches of the international diamond industry and will welcome the opportunity of service to everyone in this and related fields." Generously illustrated with photographs.

# Here and There

### New Appointments

The National Bureau of Standards recently announced three new appointments. Chester H. Page has been named Electronics Consultant, John E. White has been appointed Chief of the Electron Tube Section of the Electronics Standards Laboratory, and Curtis J. Humphreys has been named Chief of the Radiometry Section of the Atomic and Molecular Physics Division.

Harold A. Zahl has been appointed Director of Research for the Signal Corps Engineering Laboratories, Fort Monmouth, New Jersey.

Kenneth D. C. Hickman is serving both Eastman Kodak Company, Rochester, New York, and Arthur D. Little, Inc., Cambridge, Massachusetts, as consultant under arrangements recently concluded. Dr. Hickman, a native of London, England, is the inventor of the modern molecular still used for refining heat-sensitive oils.

### Necrology

William Gaertner, 84, president of the Gaertner Scientific Corporation and designer of scientific instruments, died December 3 at Wilmette, Illinois.

Frederick Gardner Cottrell, scientist and inventor, died suddenly November 16 of a heart attack while attending the National Academy of Sciences meeting at the University of California, Berkeley. He was 71. He organized Research Corporation of New York in 1912 and endowed it with the income of valuable patents.

Cecil E. Haller, physicist and radio engineer, died at the age of 40 at Lancaster, Pennsylvania, in July 1948. He was chiefly known for the development of the 829B transmitting tube.

### Division of Solid State Physics

The annual meeting of the Division of Solid State Physics, American Physical Society, will be held March 10, 11, and 12, 1949, in Cleveland, Ohio. Members should make reservations at the Hollenden Hotel, Cleveland, not later than March 1.

### Meeting of Crystallographic Society

The fourth Annual Spring Meeting of the Crystallographic Society of America will be held at the University of Michigan, Ann Arbor, Michigan, April 7, 8, and 9, 1949. Rooms and meals will be available at the Michigan Union, which will handle reservations. At least one of the sessions will be devoted to a symposium on some special topic of general crystallographic interest. For further information address Dr. Howard T. Evans, Jr., Laboratory for Insulation Research, Massachusetts Institute of Technology, Boston 39, Massachusetts.

### I.S.A. Convention Locations for 1949 and 1950

The following dates and places for future conferences and exhibits of the Instrument Society of America have been selected: September 12–16, 1949, Municipal Auditorium, St. Louis, Missouri; September 11–15, 1950, Buffalo Auditorium, Buffalo, New York.

### Symposium on Molecular Structure and Spectroscopy

A Symposium on Molecular Structure and Spectroscopy will be held at the Mendenhall Laboratory of Physics at The Ohio State University from June 13 through June 17, 1949. There will be discussions of the interpretation of molecular spectroscopic data as well as methods of obtaining such data. In addition, there will be sessions devoted to those phases of spectroscopy of current interest. A dormitory will be available for those who wish to reside on the campus during the meeting. For further information, or for a copy of the program when it becomes available, write to Professor Harald H. Nielsen, Mendenhall Laboratory of Physics, The Ohio State University, Columbus 10, Ohio.